

# 基于语义的汉语普通话电子喉语音转换增强

钱兆鹏<sup>1</sup>, 肖克晶<sup>2</sup>, 刘 蝉<sup>3</sup>, 孙 悦<sup>1</sup>

(1. 北京航空航天大学生物与医学工程学院, 北京 100191; 2. 中国人民大学信息学院, 北京 100872;  
3. 北京机械设备研究所, 北京 100854)

**摘 要:** 电子喉语音存在基频单一、发声机械、辐射噪声大等多种缺陷,这严重影响了电子喉语音可懂度和自然度,特别是对汉语普通话之类的声调语言,问题尤其严重. 汉语普通话电子喉语音识别存在辅音混淆的问题并且识别结果没有声调,因此本文在识别结果的基础之上设计了拼音拼写修正器和声调标注工具,再结合基于 Tacotron-2 的 TTS 实现了电子喉语音向正常语音的转换. 客观评价实验结果表明,拼音拼写修正器可以提高拼音准确率,声调标注在有上下文的语义环境中具有较高准确率. 主观听力测试结果表明,本文所提方法在不同语言水平上提高了汉语普通话电子喉语音的可懂度和自然度. 研究表明,本文设计的方法可以将不带声调的电子喉语音转换为正常语音,相比于传统语音转换方法具有更高的性能.

**关键词:** 电子喉语音; 拼音修正; 拼音声调标注; 语音转换

**中图分类号:** TP391.42      **文献标识码:** A      **文章编号:** 0372-2112 (2020)05-0840-06

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2020.05.002

## Voice Conversion for Enhancing Mandarin Electro-Laryngeal Speech Based on Semantic Information

QIAN Zhao-peng<sup>1</sup>, XIAO Ke-jing<sup>2</sup>, LIU Chan<sup>3</sup>, SUN Yue<sup>1</sup>

(1. School of Biological Science & Medical Engineering, Beihang University, Beijing 100191, China;  
2. Information School, Renmin University of China, Beijing 100872, China;  
3. Beijing Mechanical Equipment Institute, Beijing 100854, China)

**Abstract:** The Electro-Laryngeal (EL) speech has some drawbacks such as single fundamental frequency, mechanical sound and large radiation noise. The drawbacks affect the intelligibility and naturalness of the EL speech. Especially, the tonal language such as Mandarin EL speech would be worse understanding. In this paper, the spelling corrector for pinyin and the tone labelling tool are designed to solve the problems that Mandarin EL speech recognition has some errors in consonants and the recognition result has no tone. The result is synthesized into the healthy speech by TTS based on Tacotron-2. The objective evaluation results show that the accuracy of pinyin spelling corrector has been improved; the accuracy of tone labelling under contextual environment is very high. The subjective results shows the proposed method can improve the intelligibility and naturalness of the EL speech a lot. The results illustrate that the proposed method can convert the EL speech without tone into the healthy speech. And the proposed method performs better than the traditional method based on speech signal processing.

**Key words:** electro-laryngeal speech; pinyin spelling corrector; pinyin tone labelling; voice conversion

### 1 引言

电子喉语音是喉切除患者借助电子喉发出的语音,具有基频单一、发声机械、辐射噪声大等多种缺陷,严重影响语音的可懂度,特别是对于声调语言(如汉语

普通话),问题尤其严重<sup>[1-3]</sup>.

为了提高电子喉语音的可懂度,研究者提出了多种处理方法<sup>[4-7]</sup>. Li Wang<sup>[8]</sup>设计的电容触摸式电子喉可以较为便捷地改变电子喉的声调,能够有效提高汉语普通话电子喉语音可懂度.然而电子喉语音辐射噪

收稿日期 2019-08-05; 修回日期 2019-10-08; 责任编辑: 马兰英

基金项目: 北京市自然科学基金(No. 4194079); 北京航空航天大学虚拟现实国家重点实验室开放课题(No. VRLAB2018B06); 北京工商大学农产品质量安全追溯技术及应用国家工程实验室开放课题(No. AQT-2018-YB4)

声大、不自然的问题仍然存在. Keigo Nakamura 等<sup>[9]</sup>基于高斯混合模型(GMM),利用 STRAIGHT 提取声学特征参数基频 F0、声道谱参数(SP)和非周期性参数(AP)分别针对电子喉语音(EL-Speech)、气动电子喉语音(air EL-Speech)以及 silent 电子喉语音(silent EL-Speech)设计了三种类型的语音转换系统(VCS),其测试效果显著,可以提高电子喉语音的可懂度.

上述方法尽管在提高言语可懂度方面都取得了一些进展,但是效果有限. Qian Z 提出使用基于 WaveNet-CTC 的方法识别汉语普通话电子喉语音<sup>[10]</sup>,然而由于喉全切除术后患者借助电子喉发音时口腔气流较弱<sup>[11]</sup>,因此电子喉语的塞音及塞擦音容易出现混淆,这使得电子喉语音识别存在瓶颈问题,准确率难以继续提升.此外,电子喉语音识别结果不包含声调,这限制了汉语普通话的可懂度.汉语普通话作为一种声调语言,其语义认知受到声调影响较大.

本文是汉语普通话电子喉语音识别(ASR)的工作延续,为了解决拼音识别错误的问题,使用基于编辑距离<sup>[12,13]</sup>的拼写纠错方法,通过构建语言模型计算得到与目标词最接近词的概率<sup>[14]</sup>.此外,本文使用线性标注的方法<sup>[15,16]</sup>为拼音添加声调.最后利用文语转换技术(TTS)<sup>[17,18]</sup>将带声调拼音直接合成为清晰自然的健康语音.本文设计了客观和主观听力测试,分别对转换后语音的声调、可懂度、自然度做出评估.

## 2 基于语义的电子喉语音增强

本文在汉语普通话电子喉语音识别的基础上,基于 Tacotron-2 设计了可以直接根据拼音合成为健康人语音的 TTS.本文所设计的汉语普通话电子喉语音转换系统包括 ASR 和 TTS 两部分,整体概述如图 1 所示.其中,TTS 的输入部分包括汉语拼音容错处理和声调标注处理.

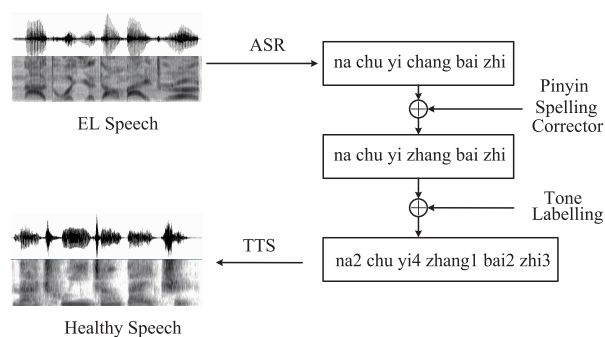


图1 汉语普通话电子喉语音转换概述

基频不变电子喉语音经过 ASR 识别为不带声调的拼音,经过拼音拼写修正器的处理,再标注声调,由 TTS 合成为带有声调的普通话语音.

### 2.1 基于 WaveNet-CTC 的汉语普通话电子喉语音识别

基于 WaveNet-CTC 技术,我们搭建了一个 ASR,用于电子喉语音识别,系统包括基于 WaveNet 的声学模型、基于 CTC 的音素解释模型,基于拼音音节的语言模型以及最大概率解码器等.汉语普通话电子喉语音识别框架详见文献<sup>[10]</sup>,汉语普通话电子喉语音识别结果为没有声调的拼音.

### 2.2 基于编辑距离的拼音修正器

本文基于编辑距离设计了拼音拼写修正器,用于修正识别结果中出现错误的拼音.拼写错误通常包括 Non-word Errors 和 Real-word Errors,前者是指词拼写本身不合法,后者是指词本身拼写合法,但在语境中存在错误.

对于 Non-word Errors,首先根据最小编辑距离计算出最相似的候选词,最小编辑距离通常涉及四种操作“插入、删除、替换、相邻字符位置变换”.其次基于噪声信道模型思想<sup>[19,20]</sup>进一步计算最大可能词的先验概率  $P(w)$  和转移概率  $P(x|w)$ ,如式(1)所示.

$$\begin{aligned} \hat{w} &= \operatorname{argmax}_{w \in W} P(w|x) \\ &= \operatorname{argmax}_{w \in W} \frac{P(w|x)P(w)}{P(x)} \\ &= \operatorname{argmax}_{w \in W} P(w|x)P(x) \end{aligned} \quad (1)$$

其中, $x$  为带有噪声的输入词, $w$  为最大可能词,式中对于特定的  $x$ , $P(x)$  不发生变化. $P(w)$  和  $P(x|w)$  可以通过训练语料库计算得到.

对于 Real-word Errors 问题,目前仍然存在很多探索<sup>[21-24]</sup>,因此本文基于 Trigram-Bayes 设计了拼音修正器.其中,Trigram 是 2 阶 HMM,可近似表示为式(2)所示.

$$P(w_3|w_1, w_2) = \frac{P(w_1, w_2, w_3)}{P(w_1, w_2)} \quad (2)$$

其中, $P(w_i)$  为第  $i$  个词的概率.利用贝叶斯计算目标词和周围词在词典中的概率,将词典中最高概率的同义词作为候选词以替代目标词及周边词,计算方式如式(3)和式(4).

$$f = \log\left(\frac{P(w_t, w_s)}{P(w_t) * P(w_s)}\right) \quad (3)$$

其中, $P(w_t)$  是目标词的概率, $P(w_s)$  是周边词概率.

$$w = \sum_{c_i \in C} f_i + \sum_{s_i \in S} \max(s_i) \quad (4)$$

其中, $\max(s_i)$  是待替换词的同义词  $s_i$  最大概率值, $f_i$  是待替换词  $c_i$  的最大概率值.

### 2.3 基于 Bi-LSTM-CNNs-CRF 的拼音声调标注

Bi-LSTM-CNNs-CRF 模型被应用于解决序列标注问题,取得了较好的效果.本文设计的拼音声调标注架构如图 2 所示.

不带声调的拼音经过 word embedding 技术得到词

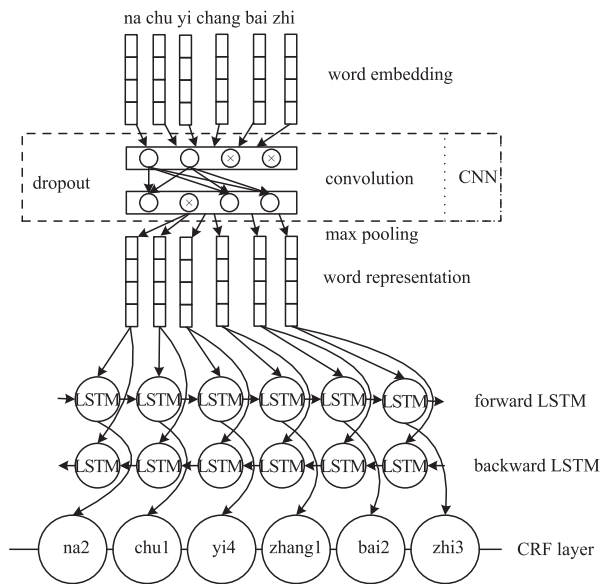


图2 基于Bi-LSTM-CNNs-CRF的拼音标注架构

向量,再由 CNN 表示为特征向量,经 Bi-LSTM 映射到输出层,CRF 根据输出层状态计算得到输入序列中每个拼音对应的声调标记.序列中所有输出标记对应的条件概率计算如式(5).

$$p(y|z; W, b) = \frac{\prod_{i=1}^n \psi_i(y_{i-1}, y_i, z)}{\sum_{y' \in Y(Z)} \prod_{i=1}^n \psi'_i(y'_{i-1}, y'_i, z)} \quad (5)$$

其中,  $\psi_i(y_{i-1}, y_i, z) = \exp(W_{y',y}^T z_i + b_{y',y})$ ,  $W_{y',y}^T$  是权重矩阵,  $b_{y',y}$  为偏置, 对应于标记对  $(y', y)$ . 使用最大条件似然估计(6)训练 CRF 模型, 训练集合由  $\{(z_i, y_i)\}$  构成.

$$L(W, b) = \sum_i \log(P(y|z; W, b)) \quad (6)$$

CRF 解码求得标记序列  $y^*$ , 由式(7)计算得到最大条件概率.

$$y^* = \operatorname{argmax}_{y \in Y(Z)} p(y|z; W, b) \quad (7)$$

训练过程和解码过程均由维特比算法<sup>[25,26]</sup>求解.

## 2.4 基于 Tacotron-2 的 TTS

当前最为先进的 TTS 技术是 Jonathan Shen 等<sup>[18]</sup>提出的 Tacotron-2, 这是一个基于注意力机制 (attention) 的端到端神经网络. Tacotron-2 包括由 Bi-LSTM 实现的编码器和解码器, 编码器用于感知输入文本信息, 解码器用于生成语音声学特征参数, 在解码器生成阶段, attention 用于关注上下文特征向量的不同部分, 设计架构详见文献[18].

## 2.5 实验设置

### 2.5.1 数据准备

本文选用标贝科技公司开放的语音合成数据集作为 Tacotron-2 的训练数据. 以基于 WaveNet-CTC 的汉语

普通话电子喉语音识别系统 (ASR) 直接对接不带拼音训练的 Tacotron-2 (TTS), 组合成 VCS 作为基线系统, 以所提出的方法作为对比系统. 标贝数据集中的汉语文本作为训练拼音修正器以及拼音声调标注工具的数据库.

### 2.5.2 评价实验设计

本文设计了评价声调变化、语音转换效果的客观评价实验以及评价语音可懂度、自然度和相似度的主观评价实验.

本研究招募 5 名志愿者, 平均年龄在 27 周岁, 听力水平正常, 作为听力受试者, 对汉语普通话电子喉语音转换后的结果做出主观评价, 可懂度和自然度评价采用平均主观得分 (Mean Opinion Score, MOS) 法对转换后语音打分. MOS 评分法分成 5 个等级, 1 - 极差, 2 - 差, 3 - 中, 4 - 良, 5 - 优.

## 3 实验结果

本文设计了客观评价测试实验用于评价转换后语音的声调准确性, 主观评价测试实验用于评价转换后语音的可懂度和自然度.

### 3.1 客观评价结果

本文基于编辑距离设计的拼音拼写修正器对识别结果做了修正, 结果如图 3 所示, 准确率统计结果如表 1 所示.

表 1 修正前后拼音准确率

拼音	准确率	标准差
Non-word 拼音修正结果	84.58%	0.083
Real-word 拼音修正结果	83.62%	0.079
原始拼音识别结果	81.56%	0.082

表 1 中的 Non-word 是指存在拼写错误的拼音, 而 Real-word 则是指符合拼写规则, 但实际意思出现错误的拼音. 这在汉语普通话电子喉语音识别结果中占比较大, 因为电子喉语音识别结果主要出现混淆的是塞音、擦音以及塞擦音.

图中实线方框是汉语普通话电子喉语音识别错误的拼音, 虚线方框是经过拼音修正器修改过的拼音. 图 3(a) 是由 ASR 的识别结果直接合成的语音基频分析, 图 3(b) 是拼音修正 + 声调标注后再合成的语音基频分析. 上述基频分析基于 Praat.

在包括字、词、短句、长句 4 个不同语言水平上, 本文统计了声调标注准确率, 如图 4 所示. 本文选择不同语言水平的测试语料数据, 目的是为了评价声调标注在不同语义难度下的性能.

从图 4 中可以得到声调标注前后合成语音在不同语言层次上的声调准确率. 无声调标注的拼音由 TTS

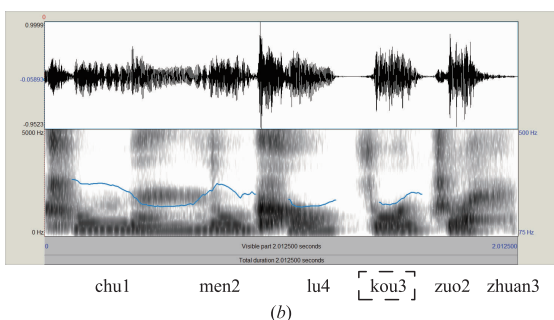
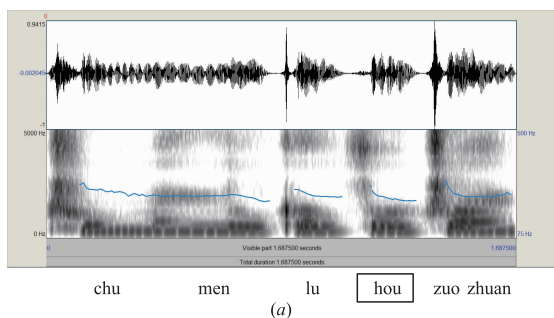


图3 转换后语音声调基频分析图

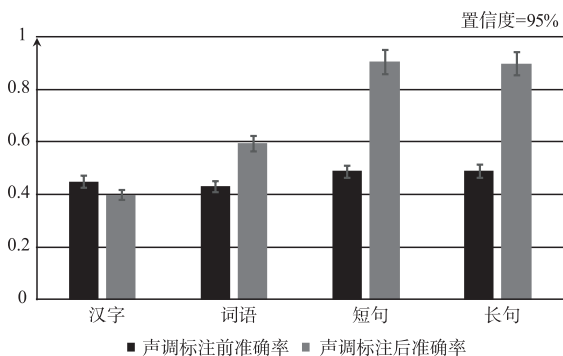


图4 声调标注前后的语音声调准确率分析

合成为语音后,整体准确率接近 50%,字和词水平上声调准确率低于短句水平上的声调准确率.声调标注后的拼音再经 TTS 合成语音后,除字水平外,词水平在 80% 以上,句子水平声调准确率在 90% 以上.

### 3.2 主观评价结果

本文所提出的方法同时适用于基频不变电子喉语音和基频可调电子喉语音.因此,我们设计了不带声调电子喉语音、带声调电子喉语音、无声调拼音直接合成语音和带声调拼音合成语音的对比试验,实验对可懂度和自然度做了分析,结果如图 5 和图 6 所示.

图 5 展示了不同语音的可懂度分析结果.其中,不带声调电子喉语音可懂度 MOS 均在 2 以下,在词语及短句语言水平上可懂度表现好于单个汉字和长句.带声调电子喉语音可懂度评价介于 2-3 之间,词语语言水平上可懂度优于其他语言水平,短句以及长句语言水平上可懂度优于单个汉字.不带声调拼音直接利用

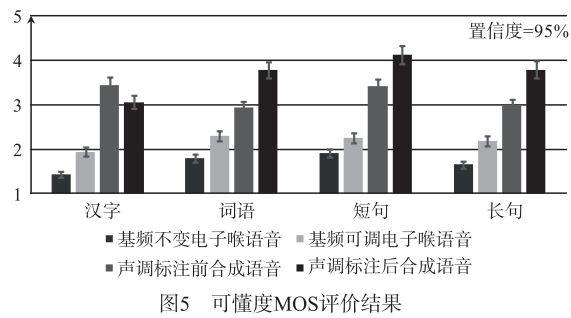


图5 可懂度MOS评价结果

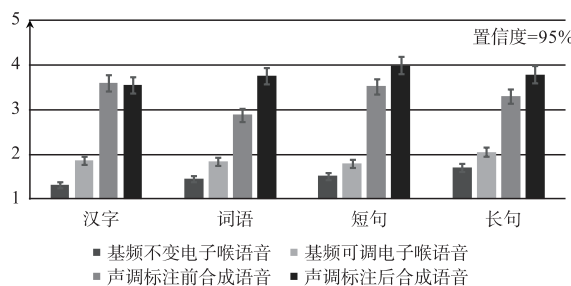


图6 自然度MOS评价结果

TTS 合成的语音可懂度 MOS 均超过 3,可懂度在各语言水平上表现均衡.带声调拼音合成的语音在词语语言水平上可懂度表现最优,MOS 值超过 4,短句及长句语言水平上表现次之,单个汉字可懂度较低.

图 6 是语音自然度的评价分析结果,不带声调的汉语普通话电子喉语音在不同语言水平上随语义信息量的增加自然度呈现递增趋势.带声调汉语普通话电子喉语音自然度表现出平稳性,在不同的语言水平上自然度 MOS 均超过 2.不带声调拼音合成语音只在单个汉字以及短句上表现出较高自然度,词语及长句自然度较低.带声调拼音合成语音均表现出较高自然度,尤其是短句语言水平上自然度最高.

## 4 讨论与展望

由于全喉切除术后患者的发音特性,汉语普通话电子喉语音的塞音、擦音以及塞擦音容易出现混淆,这是汉语普通话电子喉语音识别的瓶颈问题.为了解决这一问题,本文设计了基于编辑距离的拼音拼写纠正器,基于编辑距离的拼写纠正方法是当前拼写纠正领域最为常用的方法<sup>[12,13]</sup>.统计分析结果说明使用拼写修正器可以提高拼音准确率,其中 Non-word 类别的错误拼音可以得到较好得修正,而 Real-word 效果不尽如人意.Humboldt 认为<sup>[27]</sup>汉语词汇具有多功能性,汉语通过句法处理来实现曲折变化的语义,这说明汉语拼音相对于英语单词而言具有更高的上下文相关性.本文设计了客观评价实验旨在分析所设计的声调标注方法的有效性.所提出方法在单个汉字水平上声调标注准确率较低,这是因为缺乏上下文信息.而 TTS 直接将不

带声调的拼音合成为语音,其声调则在不同语言水平上表现出较为稳定的准确性,但准确率 50% 上下,呈现出一定的随机性. 相比之下,经过声调标注后的拼音则在词语、短句以及长句语言水平上表现出较高的准确率.

拼音的准确性与否影响着声调的标注,拼音修正和声调标注两部分工作影响最终合成语音的可懂度和自然度. 主观评价分析结果说明,在没有声调变化的情形下,语义信息量的增加对汉语普通话的理解具有促进作用. 但如果在声调变化的情形下,语义信息量增加会提高声调变化的复杂度,声调错误率提升反而会增加对汉语普通话的理解难度. 因此,声调标注准确性对于提高转换后语音的可懂度和自然度极其重要. 结合主客观评价结果来看,声调标注较高准确率对应于转换后语音的较高可懂度和自然度,在有上下文的语义环境中,本文提出的方法要优于直接将识别结果合成为语音的结果.

尽管如此,本文设计的方法仍然存在不足,例如在拼音准确率上提高有限. 在接下来的研究中,我们将进一步探索基于语义距离设计具有更高准确率的拼音拼写修正器. 除此之外,本文设计的方法计算开销较大,因此节省资源开销是一个重要的研究方向.

尽管存在上述不足,然而本文所设计的方法仍然可以极大程度地提升电子喉语音的可懂度和自然度. 而且,本文所设计的 ASR 结合 TTS 的方法相较于传统语音转换的方法能够获得与目标说话人更为接近的转换后语音. 在拼音层面上利用自然语言处理技术标注声调后,使得基频不变电子喉语音依然可以转换成带有声调的正常语音,这是传统语音转换技术所不具有的优势.

## 5 结论

本文在汉语普通话电子喉语音识别结果的基础上,基于自然语言处理技术分别设计了拼音拼写修正器和声调标注工具,再结合 TTS 将带声调的拼音合成为正常语音. 主客观评价实验结果表明,本文所提出的方法表现出更高的可懂度和自然度. 可以解决电子喉语音辐射噪声大、发声机械导致的语音可懂度和自然度较低的问题.

### 参考文献

[1] Liu H, Ng M L. Electrolarynx in voice rehabilitation [J]. *Auris Nasus Larynx*, 2007, 34(3): 327 – 332.  
 [2] Watson P J, Schlauch R S. Fundamental frequency variation with an electrolarynx improves speech understanding: A case study [J]. *American Journal of Speech-Language Pa-*

*thology*, 2009, 18(2): 162 – 167.

- [3] Guo L, Nagle K F, Heaton J T. Generating tonal distinctions in Mandarin Chinese using an electrolarynx with preprogrammed tone patterns [J]. *Speech Communication*, 2016, 78: 34 – 41.  
 [4] Goldstein E A, Heaton J T, Stepp C E, Hillman R E. Training effects on speech production using a hands-free electromyographically controlled electrolarynx [J]. *Journal of Speech, Language, and Hearing Research*, 2007, 50(2): 335 – 351.  
 [5] Wang L, Feng Y, Yang Z, Niu H J. Development and evaluation of wheel-controlled pitch-adjustable electrolarynx [J]. *Medical & Biological Engineering & Computing*, 2017, 55(8): 1463 – 1472.  
 [6] Goldstein E A, Heaton J T, Kobler J B, Stanley G B, Hillman R E. Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity [J]. *IEEE Transactions on Biomedical Engineering*, 2004, 51(2): 325 – 332.  
 [7] Liu H J, Zhao Q, Wan M X, Wang S P. Enhancement of electrolarynx speech based on auditory masking [J]. *IEEE Transactions on Biomedical Engineering*, 2006, 53(5): 865 – 874.  
 [8] Li W, Zhaopeng Q, Yijun F, Haijun Niu. Design and preliminary evaluation of electrolarynx with F0 control based on capacitive touch technology [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, 26(3): 629 – 636.  
 [9] Nakamura K, Toda T, Saruwatari H, Shikano K. Speaking-aid systems using GMM-based voice conversion for electrolaryngeal speech [J]. *Speech Communication*, 2012, 54(1): 134 – 146.  
 [10] Qian Z, Wang L, Zhang S, Liu C, Niu H. Mandarin electrolaryngeal speech recognition based on wWaveNet-CTC [J]. *Journal of Speech, Language, and Hearing Research*, 2019, 62(7): 2203 – 2212.  
 [11] Hyman M. An experimental study of the relative sound pressure, duration, intelligibility, and aesthetic aspects of the speech of artificial-larynx, esophageal, and normal speakers [D]. USA: The Ohio State University, 1953.  
 [12] Levenshtein V I. Binary codes capable of correcting deletions, insertions, and reversals [J]. *Soviet Physics Doklady*, 1966, 10(8): 707 – 710.  
 [13] Ristad E S, Yianilos P N. Learning string-edit distance [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(5): 522 – 532.  
 [14] Kukich K. Techniques for automatically correcting words in text [J]. *Acm Computing Surveys (CSUR)*, 1992, 24(4): 377 – 439.

- [15] Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging [ J/OL ]. <https://arxiv.org/abs/1508.01991v1>. 2015-08-09.
- [16] Ma X, Hovy E. End-to-end sequence labeling via bi-directional lstm-cnns-crf [ J/OL ]. <https://arxiv.org/abs/1603.01354>. 2016-05-29.
- [17] Toda T, Tokuda K. A speech parameter generation algorithm considering global variance for HMM-based speech synthesis [ J ]. IEICE TRANSACTIONS on Information and Systems, 2007, 90(5): 816-824.
- [18] Shen J, Pang RM, Weiss RJ, Schuster M, Jaitly N, Yang ZH, Chen ZF, Zhang Y, Wang YX, Skerry-Ryan RJ, Saurous RA, Agiomyriannakis Y, Wu YH. Natural tts synthesis by conditioning wavenet on mel spectrogram predictions [ A ]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [ C ]. Calgary: IEEE Press, 2018. 4779-4783.
- [19] Kernighan M D, Church K W, Gale W A. A spelling correction program based on a noisy channel model [ A ]. Proceedings of the 13<sup>th</sup> conference on Computational Linguistics-Volume 2. Association for Computational Linguistics [ C ]. Finland: ACL Press, 1990. 205-210.
- [20] Brill E, Moore R C. An improved error model for noisy channel spelling correction [ A ]. Proceedings of the 38<sup>th</sup> Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics [ C ]. Hong Kong: ACL Press, 2000. 286-293.
- [21] Golding A R. A Bayesian hybrid method for context-sensitive spelling correction [ J/OL ]. <http://arxiv.org/abs/cmp-lg/9606001v1>. 1996-01-03.
- [22] Mangu L, Brill E. Automatic rule acquisition for spelling correction [ A ]. ICML [ C ]. Nashville: ICML Press, 1997. 97: 187-194.
- [23] Fossati D, Di Eugenio B. I saw TREE trees in the park: How to Correct Real-Word Spelling Mistakes [ A ]. LREC [ C ]. Marrakech: LREC Press, 2008.
- [24] Gupta S, Sharma S. A Spelling Mistake Correction (SMC) Model for Resolving Real-Word Error [ M ]. Computational Intelligence in Data Mining—Volume 1. Springer, New Delhi, 2016. 429-438.
- [25] Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm [ J ]. IEEE transactions on Information Theory, 1967, 13(2): 260-269.
- [26] Forney G D. The viterbi algorithm [ J ]. Proceedings of the IEEE, 1973, 61(3): 268-278.
- [27] Humboldt W. On Language: The Diversity of Human Language-Structure and Its Influence on the Mental Development of Mankind [ M ]. Cambridge: Cambridge University Press, 1988.

#### 作者简介



钱兆鹏(通信作者) 男, 1989 年生于山东日照. 现为北京航空航天大学博士研究生. 主要研究方向为人工智能与言语康复.  
E-mail: qianzhaopeng@buaa.edu.cn



肖克晶 女, 1991 年生于河南信阳. 现为中国人民大学博士研究生. 主要研究方向为自然语言处理, 语义分析与文本挖掘.  
E-mail: xiaokejing0501@163.com