



Automatic detection of glottal stop in cleft palate speech



Ling He^a, Jing Zhang^{a,*}, Qi Liu^a, Junpeng Zhang^a, Heng Yin^{b,*}, Margaret Lech^c

^a School of Electrical Engineering and Information, Sichuan University, No. 24 South Section 1, 1st Ring Road, Chengdu, Sichuan, 610065, China

^b Department of Cleft Lip and Palate, Hospital of Stomatology, Sichuan University, No.14, 3rd Section of Ren Min Nan Rd, Chengdu, Sichuan, 610041, China

^c School of Electrical and Computer Engineering, RMIT University, Australia

ARTICLE INFO

Article history:

Received 19 September 2016

Received in revised form 21 March 2017

Accepted 20 July 2017

Keywords:

Cleft palate speech

Glottal stop

Articulation disorder

Compensatory error

ABSTRACT

The speech therapy is essential during the whole treatment of Cleft Palate (CP) over many years. The automatic evaluation of CP speech could provide an effective assistant diagnosis. The glottal stop is a typical compensatory articulation error in CP speech. It is produced by adducting the vocal folds as a substitution during the pronunciation of oral pressure consonants. The existence of glottal stop has a great impact on speech intelligibility. In this work, an automatic glottal stop detection system is proposed. The CP speech database is collected by the Hospital of Stomatology, Sichuan University, which has the largest number of CP patients in China. This database includes extensive CP speech samples annotated by speech-language pathologists at the phoneme level. The vocabulary of this database includes all the initial consonants in Mandarin. The automatic initials and finals segmentation method is proposed firstly. Then, for the pressure initial consonants, four types of acoustic features are extracted: Mel-frequency cepstral coefficients, formants, Gammatone filtering energy feature, wavelet packet energy and Shannon entropy features. The features are processed using a k-nearest neighbor classifier. The k-fold cross validation is used to calculate the average detection accuracy, which reaches up to 93.5% using Gammatone filtering energy feature.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Cleft Lip (CL) and Cleft Palate (CP) are the most common congenital craniofacial deformity. The occurrence rate of cleft lip and cleft palate is 1.82‰ in China [1] and around 0.2‰~2.3‰ all over the world [2,3]. Compared with cleft lip, the primary impact of cleft palate is speaking problem, which is mainly caused by the defects and deformities of palate bones and soft tissues. Cleft palate can be corrected by multiple surgeries. The treatment of cleft palate includes not only the morphological and structural reconstruction, but also the functional restore [4,5]. After the first stage of cleft palate repair surgery, there are still around 30% to 50% patients suffering from speech disorders. Speech therapy will be very beneficial to help solve CP speech problems. The consistent speech evaluation provided by the professional speech pathologists are essential during the whole treatment of cleft palate over many years. Currently, the assessment of CP speech is performed by professional speech-language pathologists. It strongly depends on their subjective judgment and experiences. The computer aided automatic

evaluation system of CP speech provides an objective and effective diagnosis to both speech pathologists and patients.

There are two types of typical CP speech performances: resonance disorder and articulation disorder [6]. The most common resonance disorder in CP speech is hypernasality, which happens at the vowels only. Several researches have been done to automatically detect the existence of hypernasality [7–9] and classify four hypernasality grades in CP speech [10,11]. Articulation disorder is another category of CP speech clinical characterization. The number of types of articulation disorders is far more than that of resonance disorders, which makes the comprehensive analysis of articulation disorders much more complicated. The typical types of misarticulation include consonant omission, consonant substitution, consonant distortion, compensatory articulation, and so on. Thereinto, the most typical compensatory errors includes glottal stop, pharyngeal fricative, pharyngeal stop, posterior nasal fricative, mid-dorsum palatal stop and so on.

The glottal stop is a typical articulation disorder among patients with cleft palate. Published literature in both English and Chinese suggest that the occurrence rate of the glottal stop in cleft patients is 60–90% [12]. The existence of glottal stop will make the speech unintelligible. Speech therapy provided by speech-language pathologists is necessary to help CP patient retrieve right pronunciation habit. Furthermore, the presence of glottal stop is a sign

* Corresponding author.

to judge whether the palatal surgery is effective. Further surgery might be required to help CP patient retrieve normal anatomic structure [12–17].

The glottal stop compensatory articulation occurs when there is inadequate intraoral air pressure for the production of pressure consonant due to the existence of VeloPharyngeal Dysfunction (VPD). The CP speakers learn to move anterior sound production position posteriorly where there is a greater amount of air pressure. The speech with the glottal stop is perceived as a brief choking or popping sound in the throat [18]. The glottal stop is typical among patients with cleft palate in other languages as well. Although there are diversities among different languages, the diagnosis, causes, and treatment of glottal stop in CP speech are the same [19–25].

Several clinicians and speech pathologists at the Hospital of Stomatology have examined the acoustic characteristics of glottal stop, mainly through the analysis of spectrograms. Those researches conducted by clinicians mainly focus on the qualitative description of the glottal stop in CP speech. Li's work [26] analyzes the Mandarin initials *k* and *t*, recorded by 20 boys with cleft palate and 20 normal peers. Their experiment results indicate that the existence of glottal stop will result in the emergence of spikes, changes of Voice Onset Time (VOT), and formant transition in the spectrograms. Zhu's work [27] shows that the VOT of glottal stop becomes shorter than that of normal speech, through the observation of spectrograms for initial consonants *t*, *s* and *d*. Chen's research [28] indicates that the formants and VOTs of CP glottal stop speech are different from that of normal speech, through the observation of spectrograms for the initials *z*, *c*, *s*, *j*, *q* and *x*.

Currently, very rare work has been done to identify glottal stop in CP speech automatically. Xiao et al. [29] investigate the automatic glottal stop detection system in Mandarin CP speech. However, that work tests one initial *d* only, using a limited size of speech samples. In Maier's work [30], MFCC is applied to detect glottal stop in German. The speech samples are recorded from 4 children. The size of tested data is limited, and it contains 32 phonemes and 31 words only. In this work, an automatic glottal stop detection method in CP speech is proposed, for all the pressure consonants in Mandarin, using an extensive size of CP speech database. This paper is organized as follows: Section 2 describes the CP speech database; Section 3 illustrates the glottal stop detection method; the experiment results are listed in Section 4; Section 5 includes the conclusions and discussions.

2. Cleft palate speech data

2.1. Cleft Palate Speech (CPS) database

The major bottleneck in the field of CP speech signal processing is the collection of CP speech database and the annotation of recordings by professional speech-language pathologists. In this work, the CP speech data are collected by the Department of Cleft Lip and Palate, Hospital of Stomatology, Sichuan University, which has the largest number of CLP patients in China. The detailed information of CPS database, including the description of participates, vocabulary list and annotation, is illustrated in Section II of our previous work [10].

2.2. Structure of a Mandarin syllable

In Mandarin, a syllable is the basic unit of a phonetic structure. One syllable contains two parts: the initial and the final. There are 21 initial consonants in Mandarin, which are composed of consonants. And there are 39 finals, which are made up of vowels or compound vowels [31].

The glottal stop occurs at the initial consonants. The initials are classified into five types: fricatives (*f*, *s*, *sh*, *r*, *x*, *h*), plosives (*b*, *p*, *d*, *t*, *g*, *k*), affricates (*z*, *c*, *zh*, *ch*, *j*, *q*), nasals (*m*, *n*) and lateral (*l*). The time duration of initial consonants are much less than that of finals. Based on the research of Mandarin phonetics [32], the plosive initials have the shortest time duration: 0–32.8 ms. While the affricate initials have the longest time duration: 0–220.3 ms. Considering the pronunciation time of initials, the speech frame length is usually chosen to be shorter than that of normal speech.

2.3. Glottal Stop Speech (GSS) database

Due to the existence of VPD, the CP speakers learn to produce sounds in a compensated way. The glottal stop compensatory articulation is produced by adducting the vocal folds as a substitution for oral pressure consonants. In Mandarin, although glottal stop is extremely common among CP patients, it is rarely observed among normal population.

The glottal stop exists at the pressure consonants only [6]. In Mandarin, there are 5 non-pressure consonants: *m*, *n*, *l*, *h* and *r*. These 5 non-pressure consonants are excluded in this work. Therefore, there are 16 out of 21 initials left [31]. In this work, the GSS database includes the Chinese syllables with these 16 initials.

The occurrence of glottal stop is intermittent in clinical practice. It means that only part of the pressure consonants is produced with the glottal stop for one CP patients. Moreover, there are CP patients who pronounce no glottal stop syllables at all.

In the collected CPS database, there are a total of 10080 Chinese words, covering various types of CP speech disorders [10]. This work studies automatic glottal stop detection method only. All the Mandarin syllables with glottal stop in the GPS database are used in this work. And approximately equal number of control samples (speech syllables without glottal stop) are also included in this GSS database. Thus there are 2371 Chinese syllables in this GSS database, including 1021 speech recordings having glottal stop and 1350 normal speech recordings. The sampling frequency is 22.05 kHz, and the annotation of glottal stop is performed at the phoneme level.

3. Automatic detection of glottal stop

The glottal stop occurs during the pronunciation of initial consonants only. This database contains a total of 2371 Chinese syllables. One syllable consists of two parts: the initial and the final. In this work, the initials and finals are segmented automatically firstly. Then, the automatic glottal stop detection method is implemented at the initials only.

3.1. Automatic segmentation of initials and finals

In Mandarin, the finals are made up of vowels or compound vowels, which are all voiced phonemes. While most of the initial consonants are unvoiced, except *m*, *n*, *l* and *r* [31]. The GSS database includes the Chinese syllables with 16 pressure initial consonants, which are all unvoiced phonemes. The following method is proposed to segment the initials and finals automatically. The proposed method locates initial/final boundary in two rounds. The rough initial/final boundary is located through Step 1 to Step 5. Then the rough boundary is refined in Step 6 to obtain more accurate boundary location. Fig. 1 illustrates an example of rough initial/final boundary location steps. The tested speech sample is recorded from a male CP patient, speaking a syllable *qu* in Mandarin, which means “go” in English. Fig. 2 illustrates the flowchart of initial/final boundary refining method.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات