Automatic detection of pharyngeal fricative in cleft palate speech

Fei He¹, Geyi Zhou¹, Xinyi He², Heng Yin³, Ling He^{1,a}

¹College of Electrical Engineering and Information Technology, Sichuan University, 610065 Chengdu, China

²West China Hospital of Stomatology, Sichuan University, 610041 Chengdu, China

³State Key Laboratory of Biotherapy, Sichuan University, 610041 Chengdu, China

Abstract. Pharyngeal fricative occurs during the production of consonants, which makes the consonants lose or weaken in cleft palate speech. In clinical application, the automatic detection of pharyngeal fricative in cleft palate speech could provide objective and effective assistant aids for speech language pathologists. In this paper, a novel acoustic parameter is proposed to detect the existence of pharyngeal fricative in cleft palate speech. This proposed acoustic feature ICPD (Independent Consonant Prominent Distribution) reflects the movement of mouth and tongue. The experimental results show that normal fricative has the higher ICPD. The extracted ICPD feature is combined with k-nearest neighbor classifier to achieve the automatic detection of pharyngeal fricative. The proposed system is tested on 127 speech samples recorded by cleft palate patients and 94 by normal speakers of controls. The overall pharyngeal fricative detection accuracy is around 90%.

1 Introduction

The communication disorder is a major problem in speakers with cleft palate. The majority of cleft palate speakers have articulation disorders. Pharyngeal fricative is a typical compensatory articulation error in cleft palate speech [1]. The existence of pharyngeal fricative makes the speech blurred and incomprehensible. Pharyngeal fricative is produced as a substitute of normal fricative by speaker with cleft palate. Pharyngeal fricative occurs during the production of consonants, which makes the consonants lose or weaken [2].

In clinical practice, pharyngeal fricative in cleft palate speech is assessed by subjective or objective methods. (1) The subjective assessments are implemented by speech language pathologists, which strongly depend on their experiences. (2) The objective assessment is implemented using medical instruments, such as ultrasonic or radiograph, to observe the motion of the phonatory organ for speakers with pharyngeal fricative. These studies use digital medical instruments to compute nasalance scores and observe changes in frequency spectrum of speech signals, to provide aided diagnosis for speech language pathologists. Garcia AF et al. [3] used nasometer 6400 (Kay PENTAX) to detect nasalance scores of patients' pronunciations. The experimental results indicate that patients with pharyngeal fricatives and hypernasality have higher nasalance scores at consonants /f/ and /s/. Zhang CH et al. [4] use CSL-4150B computer system to analyze pharyngeal fricative and glottal stop in cleft palate speech. The experiment results indicate that the main cause of pharyngeal fricative is the lack of air pressure in oral cavity, to generate plosion and friction, which are necessary for production of oral consonants. Zhu YS et al. [5] use USSA Speech Spectrum Analysis System to study

the motion of throat in the process of consonants pronunciation.

Currently, a few studies have been done to achieve an automatic detection of pharyngeal fricative through investigating the acoustic characteristics of pharyngeal fricative in cleft palate speech.

Xiao Y et al. [6,7] use one-third octave combined with Logistic regression to detect pharyngeal fricative. They have found that the speech signals with pharyngeal fricative have higher amplitude below 4,000Hz and lower amplitude above 4,000Hz than normal fricative.

In this paper, a novel acoustic parameter is proposed to detect the existence of pharyngeal fricative in cleft palate speech. This proposed acoustic feature ICPD (Independent Consonant Prominent Distribution) reflects the movement of mouth and tongue. Since the movements of articulatory organs of speakers with pharyngeal fricative are different from normal speakers. KNN (K-Nearest Neighbor) classifier is employed for classification in this work.

2 Automatic detection method of pharyngeal fricative

Figure 1 is a lateral view schematic illustration of pharyngeal fricative compensatory articulation. The dotted line in Fig.1 shows the lingual configuration and placement in the production of pharyngeal fricative [8]. Pharyngeal fricative is one type of post-uvular compensatory articulations. It is produced when the base of the tongue approaches to the posterior pharyngeal wall [9], which makes the vocal tract narrowed. The airflow from the lung goes along this narrow road. It raises the occurrence potential of pharyngeal fricative. When the

^a Corresponding author: ling.he@scu.edu.cn;

pressure from lung is not adequate, the consonant is lost or weak. Most of cleft palate patients with pharyngeal fricative suffer from weak intraoral pressure. Figure 2 and Figure 3 show the waveforms of normal fricative and speech with pharyngeal fricative, respectively. From Figure 2, it is seen that the consonant is weakened comparing to normal fricative.

Figure 1. Lateral view schematic illustration of pharyngeal fricative compensatory articulation.

Figure 2. Time domain waveform of Mandarin syllable /si/ spoken by cleft palate speaker with pharyngeal fricative.

Figure 3. Time domain waveform of Mandarin syllable /si/ spoken by controlled speaker.

In this paper, energy concentrated frequency area is used to reflect the differences in pharyngeal fricative and normal speech. Energy concentrated frequency area is sensitive to the changes of the consonant. Figure 4 and Figure 5 plot the energy concentrated frequency area of speech frames with and without pharyngeal fricative. From Figure 5, it is seen that for the normal speech, the prominent peaks of energy concentrated frequency area are obvious. The center frequency of the first peak is around 5500Hz, and the center frequency of the second peak is about 13000Hz. Moreover, for the normal speech, the amplitudes of the first and second peaks are similar. The values of these amplitudes are both over zero. From Figure 6, it is seen that for the cleft palate speech with

pharyngeal fricative, the bandwidth between two peaks becomes larger, compared with normal speech. The first peak is obviously with higher amplitude, but the second peak nearly disappears with lower amplitude.

Figure 4. Energy concentrated frequency area of normal speech frames.

Figure 5. Energy concentrated frequency area of speech frames with pharyngeal fricative.

Based on the differences of energy concentrated frequency area between normal speech and pharyngeal fricative speech, a novel algorithm based on energy concentrated frequency area is proposed. This proposed acoustic feature ICPD (Independent Consonant Prominent Distribution) calculates the prominent peak distribution of the energy concentrated frequency area. ICPD is computed by the following steps.

The vocal tract can be represented by all pole filters as follows [10]:

$$H(z) = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (1)$$

where $H(z)$ is the vocal tract transfer function in z -domain. G is the gain parameter, a_k is the coefficient of the digital filter, p is the order of the all-pole filter.

$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i} \quad (2)$$

$A(z)$ is the prediction error filter. We assume that $z_i = r_i e^{j\theta_i}$ as the Arbitrary complex root, conjugate value $z_i^* = r_i e^{-j\theta_i}$, F_i is the center frequency of energy concentrated frequency area. The bandwidth of energy concentrated frequency area is B_i . F_i and B_i are calculated by following equations:

$$\begin{aligned} F_i &= \theta_i / (2\pi T) \\ B_i &= -\ln r_i / \pi T \end{aligned} \quad (3)$$

Where, T is the sampling frequency. The order p of the

prediction error filter is preestablished. In this experiment, p is set as 4. The proposed ICPD for each speech frame is estimated as follows:

$$ICPD = \left[\frac{(\prod A_i)^a}{(\prod F_i)^b} \right]^c \quad (4)$$

A_i is the amplitude of each prominent peak of the energy concentrated frequency area. F_i is the corresponding center frequency. a , b and c are the modified factors respectively, in this experiment, $a = b = c = 1$.

3 Speech materials

In this experiment, the tested speech samples are recorded by 50 cleft palate patients with pharyngeal fricative, and 44 controls. The speech signals are sampled at a rate of 44.1KHz. A total of 221 samples (127 pharyngeal fricative, 94 normal fricative) are used for pharyngeal fricative detection in this experiment. The speech utterances are recorded in a quiet room.

4 Experiments and results

In this work, a novel acoustic feature ICPD (Independent Consonant Prominent Distribution) is proposed to detect pharyngeal fricative in cleft palate speech. Figure 6 illustrates the block diagram of proposed method. The speech signal is preprocessed firstly. The energy concentrated frequency area is calculated. Then, ICPD feature is calculated.

Figure 6. The block diagram of ICPD calculation.

Figure 7 and Figure 8 plot the values of ICPD for pharyngeal fricative and normal speech frames. It's obvious that the values of ICPD in normal speech is larger than pharyngeal fricative speech:

$$ICPD_{Normal-fricative} > ICPD_{Pharyngeal-fricative} \quad (5)$$

This can be explained by the pronunciation process of pharyngeal fricative and normal speech. The constriction places for pharyngeal fricatives are in the lower pharynx, which is more posterior than those for normal fricatives

produced in the front of the mouth. This pronunciation process leads to pharyngeal fricatives having a larger anterior cavity than normal fricatives. Longer anterior cavity results in energy concentration in the lower frequency range [6], and the energy in the higher frequency is weakened.

Figure 7. ICPD for normal speech frames

Figure 8. ICPD for pharyngeal fricative speech frames.

Figure 9 plots the ICPD values for both pharyngeal fricative and normal speech frames. It is seen that the ICPD values for all pharyngeal fricative speech frames are smaller than that of normal speech frames.

Figure 9. ICPD values for pharyngeal fricative and normal speech frames

In this experiment, a total of 221 speech samples are tested using ICPD acoustic feature combined with KNN classifier. KNN can exhibit the categorical representation of data distribution in different dimensions. ICPD is used as the acoustic feature for detection of pharyngeal fricative in cleft palate speech. Speech signal is non-stationary in nature, but it can be assumed to be stationary over short

duration called frames by windowing[11]. Hamming window is used in this experiment. The influence of frame length is discussed in this work. The pharyngeal fricative detection accuracies are listed in Table 1 and Figure 10.

Table 1. The pharyngeal fricative detection accuracies with various speech frame length

Frame length (ms)	Pharyngeal fricative detection accuracy (%)
34.01	79.77
36.28	82.35
38.55	91.00
45.35	92.99
49.88	93.09
56.68	90.90

Figure 10. The pharyngeal fricative detection accuracies with various speech frame length

Table 2 listed the confusion matrixes that pharyngeal fricatives are detect by feature ICPD with various speech frame length.

Table 2. Confusion matrixes of pharyngeal fricative detection system with various speech frame length.

confusion matrix		pharyngeal fricatives	Normal fricatives
Frame length/ms 56.68	pharyngeal fricatives	112	2
	Normal fricatives	15	92
Frame length/ms 49.88	pharyngeal fricatives	113	2
	Normal fricatives	14	92
Frame length/ms 36.28	pharyngeal fricatives	96	7
	Normal fricatives	31	87
Frame length/ms 34.01	pharyngeal fricatives	90	7
	Normal fricatives	37	87

Form Table 2, it is seen that the proposed ICPD feature achieves a good performance on the detection of

pharyngeal fricatives in cleft palate speech. The detection accuracy reaches up to 93% where the speech frame is set as 49.88ms. When the length of speech frame is shorter than 34.01ms, the detection accuracy decreases dramatically.

5.Conclusion

For the cleft palate speakers, the existance of articulation disorders influences their daily communication. Pharyngeal fricative is one of the mostly occurred articulation disorders in cleft palate speech. It makes the speech blurred and misunderstood, since pharyngeal fricative occurs during the production of consonants, which makes the consonants lose or weaken. There are acoustic differences between pharyngeal fricative and normal fricative in the frequency domain when the motion of the pronunciation organ changed. In this work, a novel acoustic feature ICPD (Independent Consonant Prominent Distribution) is proposed to detect the pharyngeal fricative in cleft palate speech. The proposed ICPD feature reflects the movement of mouth and tongue. The extracted ICPD feature is combined with k-nearest neighbor classifier to achieve the automatic detection of pharyngeal fricative. The proposed system is tested on 221 speech samples recorded by cleft palate patients and equal number of controls. The overall pharyngeal fricative detection accuracy is around 90%.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No.61503264).

REFERENCES

1. Lian Ma, *Chinese Stomatological Association*, 2(1998).
2. Liping Jiang, Guoming Wang, Yusheng Yang. *CJOMS*, **1**(2005).
3. A.F. Garcia , V.C. Marino , M.I. Pegoraro-Krook , et al. *Codas*, **26**(2014).
4. Chuhan Zhang, Haiyan Zhou, Xiaohun Jiao. *J Harbin Med Univ*, **50**(2016).
5. Yunshan Zhu, Jiejun Shi. *J Pract Stomatol*, **3**(2004).
6. Yan Xiao. *Beijing Jiaotong University*, (2016).
7. Yan Xiao, Mangui Liang. *Proceedings of the 4th International Conference on Computer Engineering and Networks*, (2015).
8. J.E. Trost, *Cleft Palate Journal*, **18**(1981).
9. J.E. Trost-Cardamone. *Diagnosis of specific cleft palate speech error patterns for planning therapy or physical management needs*. In: Bzoch KR (edit). *Communicative disorders related to cleft lip and palate*. 5.ed. Austin: Pro-Ed, (2004).
10. Zhiyong Song. *Application of MATLAB in speech signal analysis and synthesis*, (2014).

11. S. Haque, M.H. Ali, A.K.M.F. Haque, *International Workshoon Computational Intelligence*, (2017).