A Study on Drinking Judgment by the Pitch Analysis of Speech Signal

Hyung-Woo Park [#1], Seong-Geon Bae[2], Myung-Jin Bae[3]
Information and Telecommunication of Department, Soongsil University
369, Sangdo-ro Dongjak-gu, Seoul
[1] pphw@ssu.ac.kr
[2] sgbae123@empal.com
[3] mjbae@ssu.ac.kr

Abstract— Voice is one of the most important parameter to identify personal data. In this paper we introduce to build an algorithm to determine whether when one is drunken by using the pitch of voice or not. The characteristic parameters of the intoxicated speech signal were extracted by comparing the voices before and after drinking using fast pitch search analysis methods.

Keyword- Intoxicated voice signal, Intoxication judgment, Pitch of voice, Speech signal processing

I. INTRODUCTION

Most accidents caused by drunk driving are serious that they even claim the lives of the drivers. Nowadays, accidents caused by drinking from the marine and aviation flight are frequent in addition to car accidents. So, the government are making several efforts to prevent the drink and driving any craft[1].

The current ways of drinking crackdown is to directly measure the degree of the driver or navigator, among which ways there are the way by the breath meter and blood sampling assay method [1-3]. Transport, railways, ships, aircraft, etc., however, are often difficult to stop while operating for measurement in close proximity. Therefore, alcohol measurement through speech analysis can be an alternative to the two methods mentioned earlier, in order to determine the flight drinking[1-2].

The simplest way to measure the remote crafts operator for drunken is to check there voice from radios. It may be difficult to determine drunken in real time. Previous study, we proposed the method to compare the voice characteristic of drunken or not. But it has several problem that the original voice is mistaken or the navigator denies being intoxicated despite of having drunk. Therefore, we propose the speech independent way of test that can detect the drunken[1-4].Chapter 2 introduce the pitch of voice and vocal track parameter. Chapter 3 introduces our proposed study. In addition, Chapter 4 the experiments and results. And Chapter 5 conclusion.

II. VOICE FEATURES AND CHANGES IN VOICE AFTER INTOXICATION

A. Linear model of voice speech system

The linear model of voice speech system has been designed by first name Fant in 1950s. This model regards voice signals as a sound-source through the vocal track, and each component of vocal track and sound-source is independent from one another. The method uses quasi-periodic pulses for voiced sounds, white-Gaussian-noise for unvoiced sounds, and glottal sounds to complete the vocal cord model[7]. Excitation sources use pulse trains and white sound for presenting the Voiced/Unvoiced signals by switching mode in time slot. In the Voiced section, the pitch period is the most important information regarding excitation source. The vocal track can be modeled as a resonance tube, or a formant frequency[8].

In this speech system model, there are two important parameters; pitch and formant. Pitch can be explained by voiced signals. A voiced sound is produced when the air is forced to push through openings between glottis (vocal cords)[6]. The tension of vocal cords controls the vibration. If the flow of air coming out through the vocal cords is closed periodically, the excitation source is vibrating quasi-periodically as it produces voiced sounds. The time from one opening of vocal cords to their next opening is denoted as fundamental period T0, and the vibrating speed of vocal cords is denoted as fundamental frequencyF0[6-8].

B. Changes in Voice after Intoxication

When we drink, the alcohol contacts the oral cavity and esophagus first before being absorbed into the gastrointestinal tract. If we drink undiluted strong liquor like soju, whiskey or kaoliang-chiew for an extended time, they cause strong stimulus to mucous membrane that surrounds the esophagus or the top layer (epidermis) and dehydration. Further, the diuretic effect of alcohol after drinking can cause the dehydration of the larynx mucosa, which phenomenon can act to the vocal cord to change the voice. It also reduces the flexibility and elasticity of the vocal cord mucosa tissue, thus may cause incomplete opening-closure of the vocal cord when sub-glottal pressure is applied, and degrades the clarity of pronunciation due to the air leak when talking after drinking. The increase in the viscosity of the vocal cord muscle and the imperfection in the opening and closing

of the glottis exalt the phonation threshold pressure, which causes the increased lung capacity and hard breathing because the voice user needs power to initiate and maintain the phonation [3-5]. As a result, the change in the energy of voice occurs and premature termination closure of the sentence is caused during pronunciation[1].

The proposed algorithm was determined through the analysis of the voice in the same order as shown in Figure 1. When the audio signal is input to segment the voice signal analysis section divides the voiced sound and the unvoiced sound, emphasizes the periodicity of the signal to better detect the pitch period from the voiced speech. Detecting the pitch periodicity in emphasis signal, and determined using the detected pitch. It does is determined by estimating the anesthesia of dehydrated alcohol, or vocal cords of the speaker's vocal cords to measure the degree of pitch change. In chapter II-A., If, as a linear prediction analysis of the speech signal to separate speech model and the excitation signal and a vocal-track parameter, where the characteristics of the analyzed source.

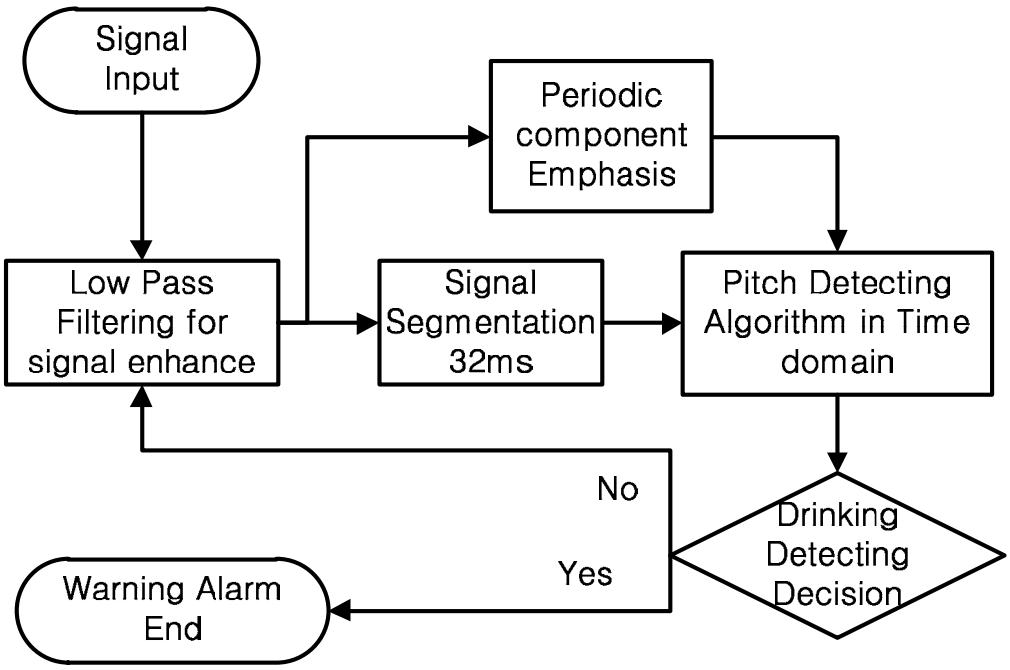

Fig. 1. Block diagram of drinking test algorithm

III. EXPERIMENTAL RESULT

Samples were taken evenly from each age group between 20's and 60's. 15 standard samples were used. Same sentence was phonated and a male voice sample without a dialect was selected. The words used for voice sample manuscript were selected from the words difficult to pronounce.

Voice samples were prepared by recording the voice of each person before and after drinking, using digital recorder. And we control the noise condition for take the clear voice from tester. And the data record by 8kHz sampling rate and 16bit quantization. In addition, the degree of alcohol was measured using a Breathalyzer before recording the same person was recorded multiple times to improve the accuracy. The timing of the recording after drinking was determined by subjective feelings and general statistical data, which is generally 2 hours after drinking 120% of one's drinking capacity.

The test data are edited using signal processing program with PC. The test sentences and word pronounced by the same person are edited in constant time intervals and rules. Matlab and C programs are used as the simulation program for the development of an algorithm to extract the parameters that determine the intoxication.

Figure 2 and Figure 3 is an experiment result of pitch contour of before and after drinking which is the plot of the same person in the same sentence. As shown in Figure 2, the pitch appears evenly before the alcohol consumption, but in Figure 3, the appearance of pitch is error and significantly uneven after ingestion.

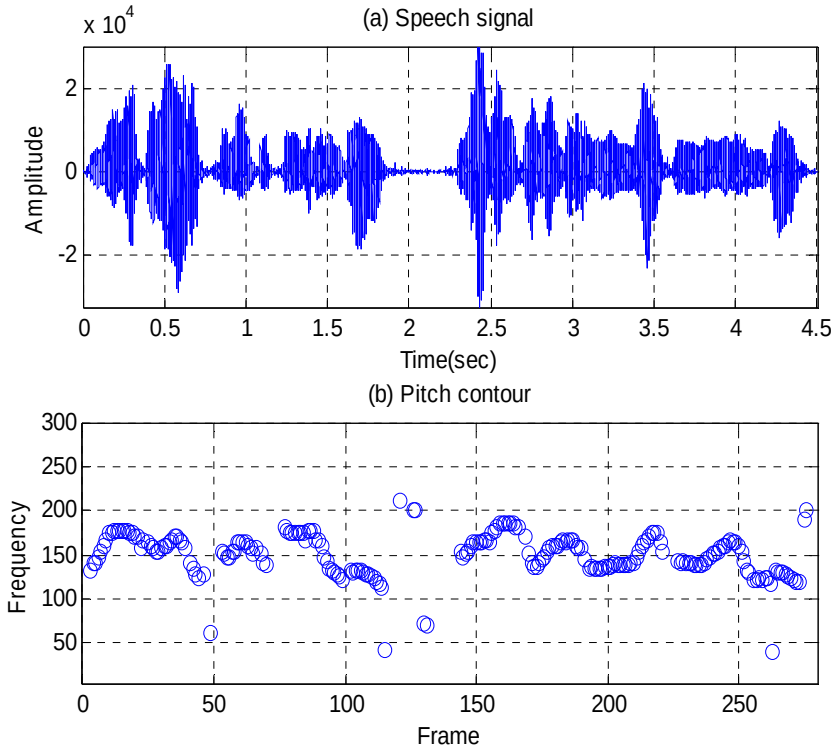

Fig. 2. Pitch contour of before drinking

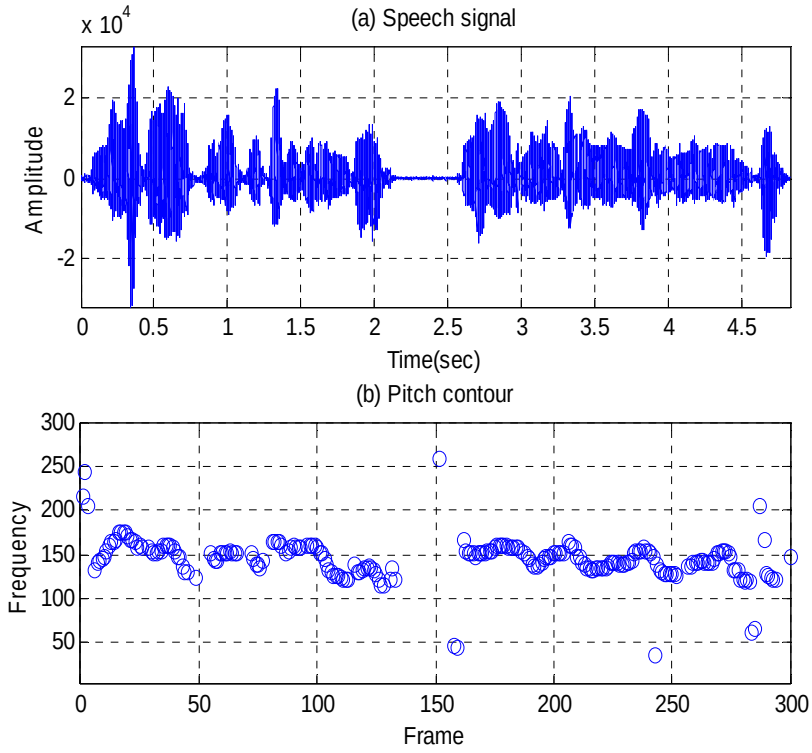

Fig. 3. Pitch contour of after drinking

Table 1 shows the pitch analysis data for four sentences and four people. In table 1, the pitch and pitch deviation are changed after drinking. And using this parameter of signal, we can determine the alcohol with a speaker-independent. Pitch change of drinkers and non-drinkers and also, also changes the distribution of the pitch was determined to be a parameter.

TABLE I. Relationship of Drinking and Pitch

	drink before		drink after	
	pitch	pitch deviation	pitch	pitch deviation
tester 1	135	42	144	60
tester 2	116	17	130	38
tester 3	149	25	148	40
tester 4	115	26	137	42
tester 5	118	37	118	43
tester 6	127	12	134	39
tester 7	131	27	127	23
tester 8	141	24	137	23
tester 9	126	32	133	30
tester 10	160	35	157	49
tester 11	145	71	134	64
tester 12	163	57	164	55
tester 13	145	55	125	14
tester 14	125	21	136	24
tester 15	119	24	125	52
average	134	34	137	40

IV. CONCLUSION

In this paper we introduce that the difference of voice when person were drunk. Especially, The pitch of voice is changed and we make a decision parameter with pitch-perturbation. The algorithm, in this paper, is using voice which is drunken and normal. Propose algorithm is weighted depending on the contribution of the accuracy of judgment on drinking. As the result of our method to 15 standard males speaker with 5% of alcohol concentration, the success rate of intoxication judgment turned out to be 80% for the same person when his voice signal was identified.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support from R&D program through Future Marine Industrial Technology Development of Korea Institute of Marine Science & Technology(grant number 20140191).

REFERENCES

[1] Seong-Geon Bae, Hyoung-woo Park, Geumran Baek and Myung-jin Bae, "A Study on Drinking Judgment Method of Speech Signal using the Ratio of Harmonic Deviation," ASK, procedding of 29th speech signal processing , 2012.
[2] Geumran Baek and Myungjin Bae, "A Study on Voice Sobriety Test Algorithm in a Time-Frequency Domain," International Journal of Multimedia and Ubiquitous Engineering, Vol.8, No.5, pp.395-402, 2013.
[3] Korean Alcohol Research Foundation, Alcohol Dictonary, Korean Alcohol Research Foundation, 2002.
[4] Bo-myung Whang And 3 others, “Analyze the acoustic characteristics of the voice change after drinking and singing,” Speech Sciences, vol.8, no 2, pp.2-4, 2001.
[5] Si-woo Lee, “A Study formant changes in drinking and non-drinking status”, Journal of the Korean Society for Academic Technology, vol.10, no.4, pp.805-810,2009. A Study on Drinking Judgment Algorithm Using Speech Parameter
[6] Hyung-Woo Park, Myung-Sook Kim and Myung-Jin Bae, "Improving Pitch Detection through Emphasized Harmonics in Time-Domain," Springer-Verlag, CCIS 352, pp. 184–189, 2012.
[7] M.J. Bae and S.H. Lee, Digital Voice Signal Analysis, Books Publishing Dong Young, 1998.
[8] Rabiner, L.R., Schafer, R.W. , Introduction to Digital Speech Processing, The Essence of Knowledge, 2007.

AUTHOR PROFILE

Hyung-Woo Park He received the Ph.D. degree in Telecommunication & Information from Soongsil University in 2013. He is currently the Professor of the Dept. of Telecommunication & Information at Soongsil University.

Seong-Geon Bae He received the Ph.D. degree in Telecommunication & Information from Soongsil University in 2014. He is currently the Professor of the Dept. of Broadcasting Sound & Visual at Daelim University.

Myung-Jin Bae He received the Ph.D. degree in Electronic Engineering from Seoul National University in 1987. He is currently the Professor of the Dept. of Information & Telecommunication at Soongsil University. He has authored and coauthored more than 300 journal articles and conference papers. His research interests are Speech Signal Processing, Audio Processing and Speech Communications.