Reducing Errors of Judgment of Intoxication in Overloaded Speech Signal

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

***Abstract*—This research looks into parameters for reducing errors of judgment intoxication by comparing speech characteristics before and after drinking using speech analysis. In this study a drinking is tested to speech signal under overloading condition, there are many of some errors in unstable conditions. Therefore, it should be determined by using judgment algorithm that is not affected by overloading. So we has searched valid frames using energy of speech before and after drinking were extracted and drinking possibility was identified by using deviation from the distance of frames. In terms of speech frame energy its value changes greatly in overloaded but energy ratio largely stays the same in normal. Therefore, after applying overloaded signals to proposed method, we get this similar result with a normal input signal was obtained.**

Keyword - Judgement of intoxication, Overloaded signal, Valid frame, Formant, pitch

I. INTRODUCTION

When it comes to social life drinking is something unavoidable. Sometimes, people recover their relationship through drinking and on the other hand, it could be something harmful. Also, people drink to have a better team work but if too much this may cause even worse relationship in the team. That is, drinking has both positive and negative effects. A few years ago, breath testing was done to the drivers on the road. However, recently, there are many different methods such as smart phone application, GPS, or one cannot start the engine after drinking by the measurer installed in the handle [2][9]. These methods are short distance measuring as measurers have to be located closer to the drinker. Different from this, if drinking is determined by speech signal it can be done even from a far distance [8]. Especially for transportations under operation such as railway, ship, airplane, and automobiles it is not easy to check but this enables their alcohol check through speech analysis. In this case, another point that we should consider is when input signal is overloaded [1]. The overload means over the limit of normal level that a certain equipment or device can handle. It may cause distortion of signal in processing [3]. Overloading happens when input microphone is placed too close, input signal level is too high or higher in time domain. Also drinking possibility is determined by using speech signal in this overloaded condition, errors will occur. Therefore, it should be measured by using method that is not affected by overloading [4][5].

As a result this research examined reducing errors using deviation of valid frames in overloaded speech signal. In session 2, it looks into checking method using deviation of valid frame. We discusses about information on experiment and result in session 3. And Session 4 provides overall conclusion of this research.

II. CHARACTERISTICS OF OVERLOADED SIGNAL

Speech signal before drinking can deliver more accurate information compared to the one after in spectrum, we used to compare distance of two frames in frequency. However, when drinking, if pressure is applied to vocal cord it may cause difficulties in complete opening and closing and thus, minimum pressure required for pronunciation is increased and lung activities rise [7]. Another is that after drinking vocal cord lacks water and makes it harder to make speech signal compared to before drinking [6][8]. Therefore, speech after drinking is delivered less accurately compared to before. We used to search frames that have higher than average energy in each frame of speech signal, valid frames can be extracted [6]. This is a valid frame.

The valid frame method used in this research focuses on the point that it has longer duration of speech signal with higher amplitude after drinking. Therefore, it can be said that more valid frames are made after drinking [5]. As non-vocalization rate increases after drinking it means it becomes harder to deliver information using speech after drinking. Here, vocalization duration becomes longer and thus, deviations of valid frames become larger after drinking. Drinking identification by using deviation of valid frames is done by applying deviation of valid frames to speech signal before and after drinking. The below equation 1 is for calculating average energy of each frame which is the reference value for imputing valid frame in speech signal.

$$E_{VF} = \frac{1}{M_a} \sum_{n=-\infty}^{\infty} \sum_{m=1}^{k} s^2(n+m) \tag{1}$$

Where M_a is number of valid frames.

In equation 1 E_{VF} is used by a reference value in valid frame and means an average in total. k is number of samples and m is number of overlapping samples in speech signal. That is, after obtaining average of values calculated from energy of each frame this is then becomes the reference value for deciding valid frame.

Based on the reference value of valid frame, valid frames are extracted. Afterwards, speech signals before and after drinking is compared by using reference deviation of values of valid frames. Figure 1 shows simple block diagram of checking drinking by using deviation of valid frame.

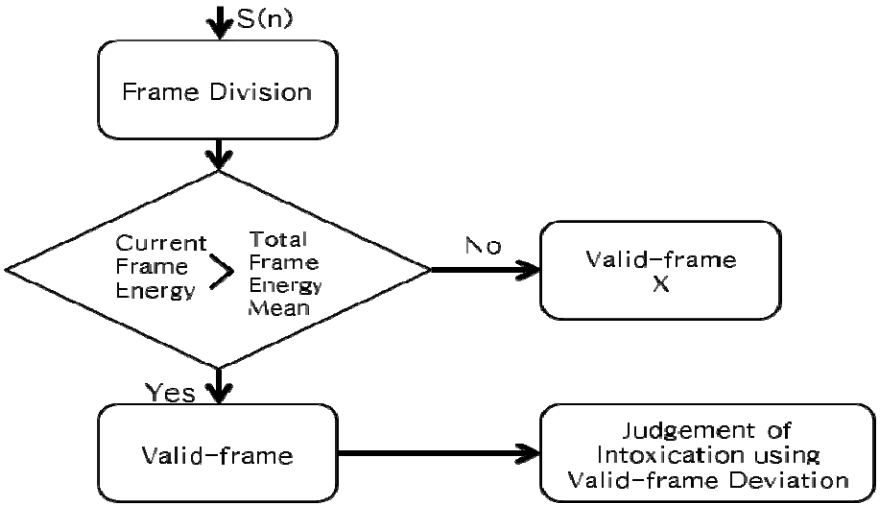
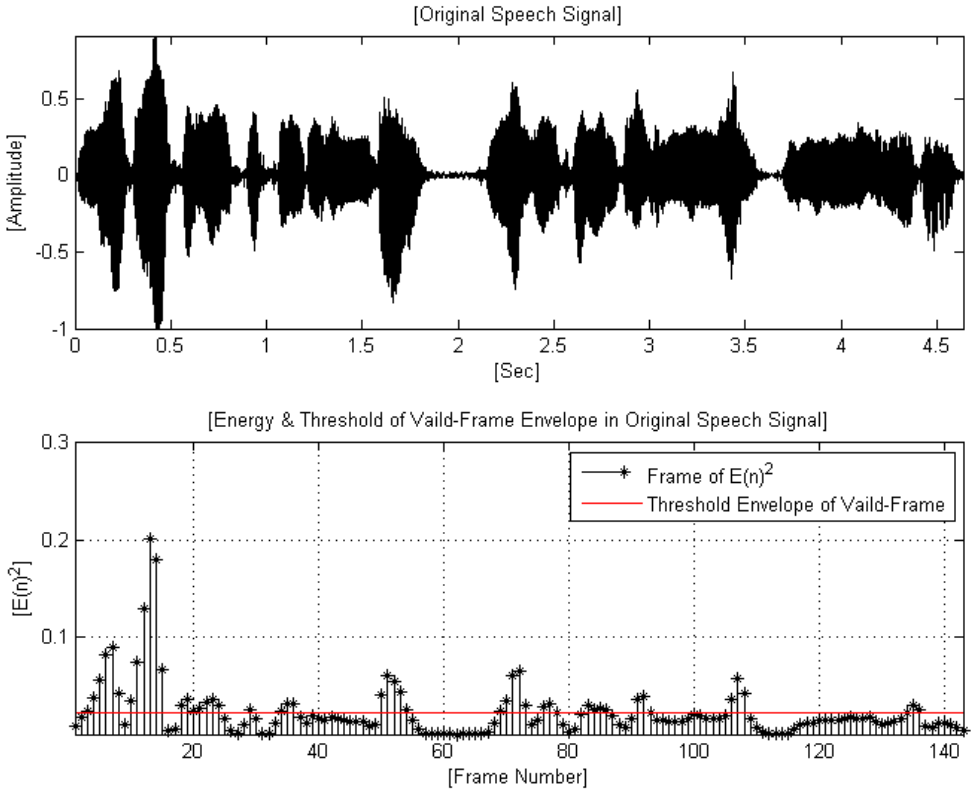


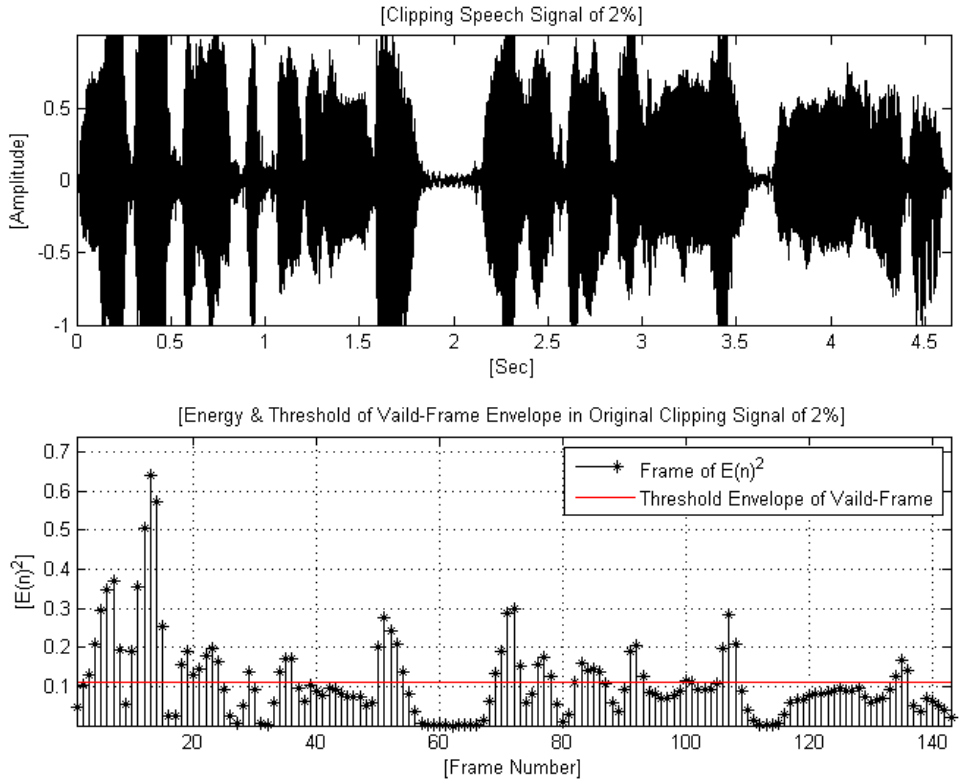
Figure 1. A judgment intoxication using a valid-frame deviation

Comparing speeches of before and after drinking, it is found that formant energy radically changes. It happens because there is a radical change in muscles at biological structure of vocal tract, making formant. An irregular relaxation and contraction of muscles at the vocal track make a sound pressure energy and it makes this biological analysis. Comparing energy with continuity of speech signal at time domain shows this characteristic. Analyzing the energy distribution for per frame shows that there is a sharp change in energy after drinking. To measure more accurate amount of this change, an analysis is done at frequency domain. The energy for each frame is calculated after comparing changes in spectrum. Figure 2 shows LPC(linear prediction coefficient) analysis method, using formant comparison, used in this paper. The method makes the amount of change in formant spectrum, compares the spectrum and applies it to the analysis of before and after drinking. Before drinking, the energy, showing the amount of change in formant, is loose and even. It makes a speech nasalized after drinking.

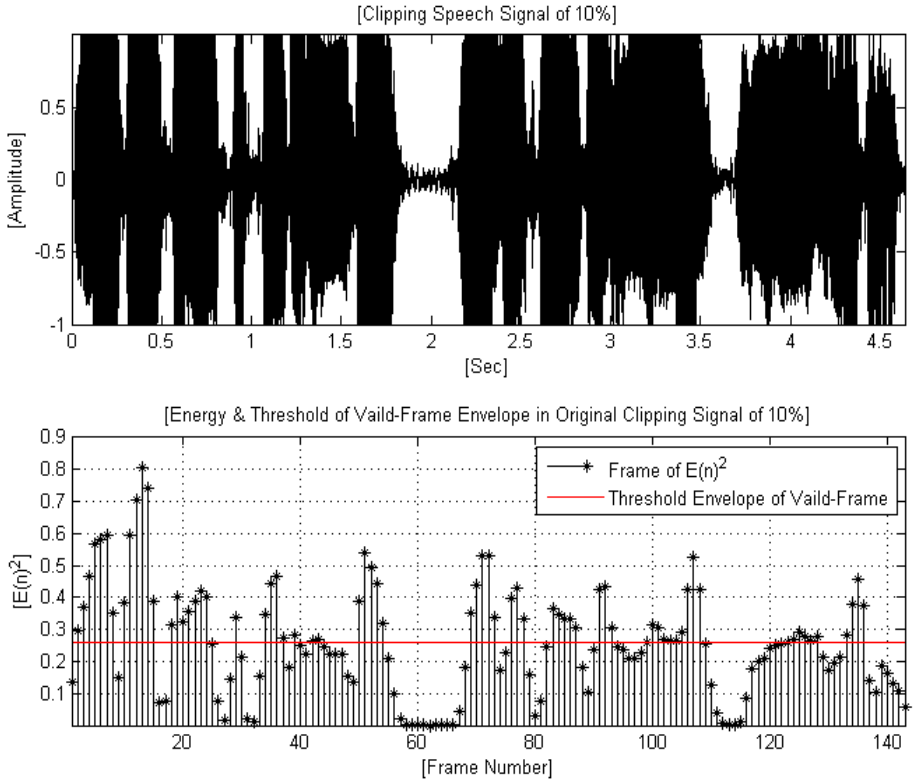
When overloaded, also speech energy increases instantly. However, even if signal is distorted due to the top and bottom waveform caused by overloading, overall outline of energy stays the same. The Figure 2 shows comparison of threshold value which determines overall energy graph and valid frame by using these characteristics. (a) shown in Figure 2 is result from input of normal speech signal and the top is overall waveform of speech signal. The second graph shows threshold value as well as reference value that determine valid frames in overall energy and one straight line. The frames that have energy going over one straight line can be referred to as valid frame. Similar to this, (b), (c) and (d) of Figure 2 shows that overall waveform of speech signal and energy has threshold value of valid frames when ratio of waveform overloading is 2% and 5% in overloaded condition.



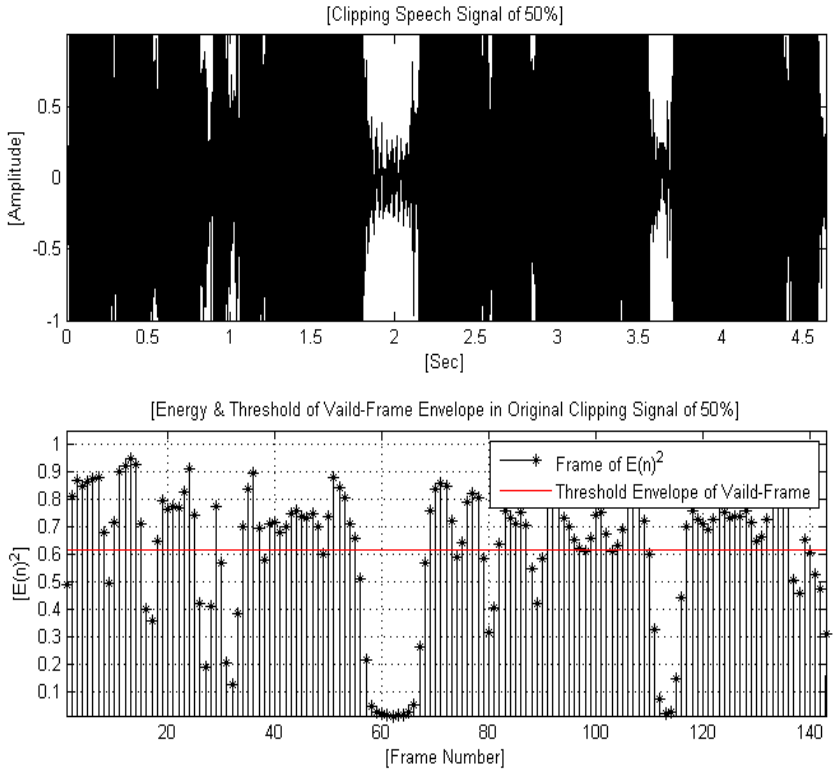
(a) When normal, graphs of totally energy deviations in valid frames.



(b) When 2% overloaded, graphs of totally energy deviations in valid frames.



(c) When 10% overloaded, graphs of totally energy deviations in valid frames.



(d) When 50% overloaded, graphs of totally energy deviations in valid frames.

Figure 2. Graphs of totally energy deviations in valid frames.

Table 1. Compared judgment intoxication before and after overloading using proposed valid frame deviation

	Compared Judgment Intoxication Before with After in Valid-Frame Deviation											
	Original Signal		2% overloaded		5% overloaded		10% overloaded		30% overloaded		50% overloaded	
	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
1	52	58	50	58	51	59	53	59	54	59	55	59
2	44	47	45	47	44	47	44	48	45	49	46	49
3	48	46	49	46	48	46	49	46	50	47	51	48
4	43	44	43	44	43	44	42	44	43	44	43	44
5	66	46	67	45	67	45	67	45	68	45	68	45
6	52	64	54	64	56	64	55	64	56	64	57	64
7	57	46	57	47	58	48	57	48	57	50	57	50
8	50	35	50	41	51	45	50	46	51	46	52	46
9	40	54	42	52	43	52	45	54	46	54	47	54
10	37	46	36	45	36	46	37	47	43	50	44	50
11	52	46	53	46	52	46	53	46	53	45	53	46
12	48	43	47	44	48	43	49	44	49	44	51	43
13	44	52	45	52	44	51	46	50	47	49	46	48
14	54	66	55	65	55	64	55	64	55	63	56	63
15	52	55	52	55	51	56	51	57	52	57	52	58
16	57	64	57	65	56	66	56	65	54	70	55	70
17	36	49	37	49	36	50	38	50	37	50	38	51
18	51	62	51	62	51	62	51	62	51	62	53	62
19	87	68	88	68	89	70	89	70	89	72	88	73
20	56	60	57	60	57	62	58	63	58	61	58	61
21	55	69	56	69	55	70	56	70	56	70	55	69
22	61	101	61	101	61	100	61	100	60	97	61	97
Avs.	52	53	51	52	51	51	52	52	52	52	53	53

Looking closely into the Figure 2, energy of the second graph increases as overloading rate rises, however, outline of the graph generally stays the same. Also, threshold value of valid frames rises along with energy, leaving similar result for number and position of valid frames. Therefore, even under the environment of overloaded signals it could be said that relatively less errors occur in this method of determining drinking by using deviation of valid energy frames.

III. CONCLUSION

Unfortunately, drink and drive check measuring device should be located close to the driver in test. Because of this reason there are many problems related to human rights and accidents. Thus, there should be a method for measuring alcohol level from far distance. One way to do so is checking drinking via driver's voice. Moreover, overloaded signal such as placing microphone too close it would be even harder to identify. So we proposed valid frame method to solve the problem, this research has applied drinking check using deviation of valid frames in frame energy. As a result, we were proven that as energy increases due to overloading threshold value of valid frames rise as well and thus, it can be said that number and position of valid frames in speech signal do not change greatly. Therefore we get the results of reducing errors of judgment intoxication in test. In the future it is expected that this can be used as a reference for checking drinking by applying not only to overloading but also to normal signals that are affected by other environmental factors.

REFERENCES

[1] Won-Hee Lee, Seong-Geon Bae and Myung-Jin Bae, "A Study on Improving the Overloaded Speech Waveform to Distinguish Alcohol Intoxication using Spectral Compensation," IJET Vol.7, No.5(2015), pp.1957-1964.
[2] Changjoon Jung, Seonggeon Bae, Myungsook Kim and Myungjin Bae, "Speech Sobriety Test Based on Formant Energy Distribution," IJMUE Vol.8, No.6(2013), pp.209-216.
[3] Seonggeon Bae, Myungjin Bae, "A New Speech Coding using Harmonics Emphasis Filter," ISAAC 2013, AACL Vol. 1(2013), pp43-44.
[4] Seonggeon Bae, Myungsook Kim, and Myungjin Bae, "On Enhancement Signal Using Non-uniform Sampling in Clipped Signals for LTE Smart Phones ", IEEE ICCE-berlin(2013), pp.125 - 126.
[5] Seonggeon Bae, Hyungwoo Park and Myungjin Bae, "On a New Enhancement of speech Signal using Nonuniform Sampling and Post Filter,"ICHIT 2012, LNCS 7425(2012), pp.723-729, Springer-Verlag.
[6] Geumran Baek, Myungjin Bae, "Study on the Judgment of Intoxication State using Speech Analysis ". Vol.352 CCIS, 2012, pp. 277-282, Springer-Verlag. 2012.
[7] Marcieli Bellé, Sílvia do Amaral Sartori, Angela Garcia Rossi. "Alcoholism: effects on the cochleo-vestibular apparatus". Rev Bras Otorrinolaringol. 73(1):116-22. 2007.
[8] Tahwinder Upile1, Fabian Sipaul1, Waseem Jerjes1, "The acute effects of alcohol on auditory thresholds, BMC Ear, Nose and Throat Disorders". 7: 4 doi : 10.1186 /1472–6815–7-4. 2007.
[9] Behne, D. M., & Rivera, S. M. (1990). Effects of alcohol on speech: Acoustic analysis of spondees. Research on Speech Perception, Progress Report, 16 (pp. 263–291). Bloomington, IN: Speech Research Laboratory,

AUTHOR PROFILE

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.