The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)
Vol. 17, No. 3, pp.47-53, Jun. 30, 2017. pISSN 2289-0238, eISSN 2289-0246

https://doi.org/10.7236/JIIBC.2017.17.3.47
JIIBC 2017-3-5

무선통신 환경에서 음성파형 보상을 통한 음주가능성 여부에 관한 연구

A Study on the Possibility of Drinking through speech Waveform Compensation in Wireless Communication Environments

이원희[*], 박형우[**], 배성근[***], 배명진[****]

Won-Hee Lee[*], Hyungwoo Park[**], Seong-Geon Bae[***] and Myung-Jin Bae[****]

요 약 해상에서는 도로가 아닌 해상 위라는 환경 때문에 음주단속을 실시하여 음주운항을 예방하기엔 어려움이 존재한다. 전 연구에서 음주여부에 따라 변화된 음성을 파악하기 위해 알고리즘을 개발되었으며, 개발된 알고리즘을 이용해 원거리의 있는 선박 운전자 그리고 선원들의 음주 가능성 여부를 파악하는 것이 가능해졌다. 음주 단속이 근거리가 아닌 원거리로 정보를 전송할 수 있는 음성을 통해 이루어진다면 거리가 얼마나 멀리 떨어져 있더라도 실시간으로 측정할 수 있다. VTS 무선 교신 환경을 이용하여 음성을 통신할 때도 무선 환경이 고르지 못할 경우에 클리핑이 일어나 음주가능성 판단율이 저하될 수 있다. 따라서 본 논문에서는 음성신호가 왜곡이 되어 음주 가능성 여부의 판단율 오차를 줄이기 위해 신호를 보상하는 방법을 제안하였다.

Abstract There is a difficulty in preventing drunken driving by enforcing alcohol control on the sea due to the environment of Marine transportation rather than roads. In the previous study, we proposed the algorithm, that was developed to identify the voices changed according to be drunk. Using the developed algorithm, it became possible to know the possibility of drinking by long distance ship operators and crew members. In that method drinking can be measured in real time, no matter how far the distance is, if the interception is through a voice that can be transmitted over a distance, rather than a short distance. When communicating voice using the VTS wireless devices, clipping occurs when that environment is uneven, and the rate of judgment of the possibility of drinking may be lowered. Therefore, in this paper, we proposed an enhanced method to compensate the signal in order to reduce the error rate of the possibility of drinking due to distortion of the speech signal.

Key Words : Drinking, Waveform-compensation, VTS, Clipped waveform, Drunken speech.

Ⅰ. Introduction

In 2015, the number of fishing vessels caught in the sea was 130 (about 68.7%), which is the highest rate.[1] Alcohol control on the road is easy to control the drinking through a short distance.[2] However, since it is necessary to stop the movement of vessels such as fishing vessels and cargo vessels on the sea, it is

*준회원, 숭실대학교 정보통신공학과
**정회원, 숭실대학교 전자정보공학부
***정회원, 강남대학교 미디어정보공학과
****정회원, 숭실대학교 전자정보공학부
접수일자: 2017년 2월 14일, 수정완료: 2017년 5월 14일
게재확정일자: 2017년 6월 9일

Received: 14 February, 2017 / Revised: 14 May, 2017 /
Accepted: 9 June, 2017
**Corresponding Author: pphw@ssu.ac.kr
Dept. of IT, Soongsil Univesity, Korea

difficult to regulate drinking through near places. On the contrary, it can be applied not only to the road but also to the maritime and air traffic if it is possible to measure the disturbance of drinking even at a remote place rather than a close range. The information that can be obtained from a distance through the state or change of the body is speech. If alcohol is interrupted through speech that can send information to a long distance, it is possible to measure any distance. In order to enforce alcohol abuse by speech over the sea, a wireless device must be provided to connect the sender and the receiver.[3-5] At sea, all vessels are managed by VTS (Sea Traffic Control System) and provide information service to help the ship to do its best. When the speech communication is performed through the VTS, the characteristics are different according to the wireless equipment, but the speech signal is distorted in the process of demodulating through the receiving equipment. In the case of a cargo ship or a large ship, it is possible to obtain a distorted speech signal by using a specially designed radio equipment, while using a relatively low specification equipment in case of a small fishing boat. When a signal is demodulated at the time of reception, a phenomenon occurs in which the distortion of the speech signal occurs during the detection process and becomes a limit. It is similar to when the input speech signal is out of the input level. If the possibility of drinking through the distorted speech signal is measured, there is an error in the judgment rate and it is difficult to control the speed.[5-8]

Therefore, in this paper, we attempt to improve the rate of decision on the possibility of drinking by compensating for distorted speech signal waveform through VTS speech communication. In Section 2, we discuss the phenomenon of speech signal distortion, and in Section 3, we discuss how to compensate the speech signal in the VTS speech communication environment proposed in this paper. In Chapter 4, we conclude the future research direction.

II. Existing speech waveform compensation method

When the speech signal is distorted, there are various cases such as deterioration due to ambient noise when the input signal is out of the input level. If the input signal is out of the input level, it is a clipped phenomenon and means that the device or device has exceeded the normal value that can be handled. Clipping may occur depending on the input amplitude level of the input microphone, and clipping may occur when the wireless environment is uneven even when communicating speech using the wireless environment. When the speech signal is clipped and distorted, harmonic distortion occurs. Figure 1 shows the spectrum with 300 Hz pure tone and 30% clipped waveform. As shown in Figure 1, when the waveform is overloaded like a square pulse, clipping phenomenon occurs, harmonic distortion occurs, and harmonic distortion of the radix occurs.[8][10-13]

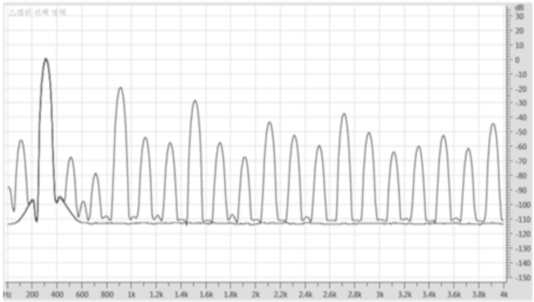


그림 1. 320Hz 순음과 35% 클리핑된 320Hz 순음 파형에 대한 주파수 특성

Fig. 1. 320 Hz with pure tone and 35% clipped waveform spectrum

To compensate for the clipped and distorted speech signal, the characteristics of the speech signal must be considered. The sound pressure energy of unvoiced sound is small and has many irregular components. The energy of a voiced sound consists of large, regular periodic components. It is necessary to distinguish the voiced and unvoiced sounds in the distorted speech signal by analyzing these characteristics. Figure 2 is a block diagram of the existing waveform compensation method. As Figure 2 shows an existing speech signal waveform compensation method it will be divided into low-frequency and high-frequency treatment of two processes. The first step is to normalize the overloaded input signal by 50%. This is because low-frequency components and high-frequency components are compensated for and then synthesized. Pass a low-pass filter of about 500 Hz to compensate for the next low-frequency part. As shown in Figure 1, overloading causes clipping and harmonic distortion. In order to compensate for the clipped signal, it is necessary to make it with the bones and the peaks, so it also functions to smoothing a square pulse-like shape using a low-pass filter. To compensate for the clipped signal, it has to be made of bones and peaks, so it uses a low-pass filter to smoothen the shape similar to a square pulse. The next step is to use a bandpass filter with a cutoff frequency of about 1000 Hz to 3500 Hz. It is necessary to compensate for the high frequency part of the speech signal. Since the harmonic distortion occurs when overload occurs, the band pass filter is used because the energy of the high frequency component of about 3500 Hz or more also increases. Then synthesize these two signals. Finally, the two signals are combined and re-synthesized using only about 50% of the signal passed through the previous band-pass filter to compensate for the high-frequency components.[8][10-13]

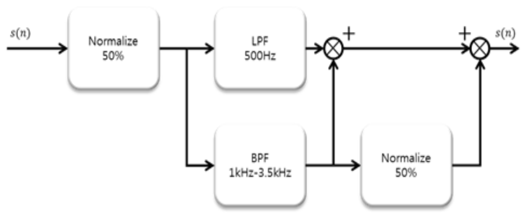


그림 2. 기존의 제안된 음성파형 보상방법[11]
Fig. 2. Block diagram existing overload speech signal compensation[11]

Ⅲ. Proposed Method and Experimental Results

When speech communication is performed using wireless communication through the VTS, if the wireless environment is uneven, distortion occurs in the speech signal. When the standby state or the weather is poor, a speech signal with deteriorated noise can be received. If the wireless communication equipment is not good, a phenomenon occurs in which not only noise but also signal clipping occurs. Figure 3 shows an example of several speech signals received at the receiving end during VTS speech communication.

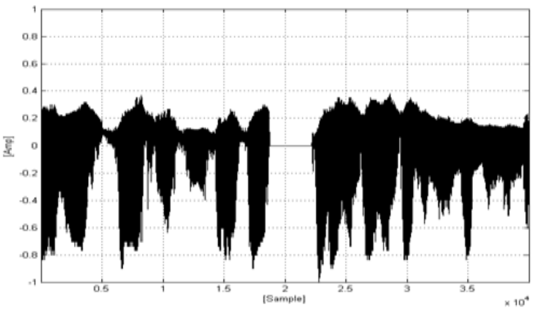


그림 3. 무선통신환경에서 음성 통신 시 수신측에서 받은 음주 전 음성신호 파형
Fig. 3. Speech signal waveform before drinking received from the receiver during speech communication in wireless communication Environments

As shown in Figure 3, it can be seen that the signal is clipped in the (+) region of the speech signal, which is an anode, as a limit. Existing waveform compensation method of speech signal is compensated similar to characteristics of original low frequency and high frequency areas when the ratio of clipping, which is a distortion rate, is about 20% or less. However, if the ratio exceeds 20%, the low frequency characteristics below 500Hz can be compensated similar to the original waveform, but compensation is difficult because the fundamental information for compensating the high frequency region is almost lost. Therefore, it is necessary to improve the high frequency domain compensation method in the existing method in order to compensate the high frequency domain. Especially, since the rate of change of the high frequency region is also used to determine the possibility of drinking, it is necessary to make a detailed compensation. The proposed method of compensating for clipped and distorted speech signals is divided into three processes: low frequency, mid frequency, and high frequency. As described in Section 2, low frequency and mid frequency components constitute the vowel portion. High frequency components are the main cause of the formation of peaks and valleys. Especially, voiced sounds are most affected by overload. Figure 4 shows a block diagram of the method for compensating the clipped speech signal proposed in this paper.

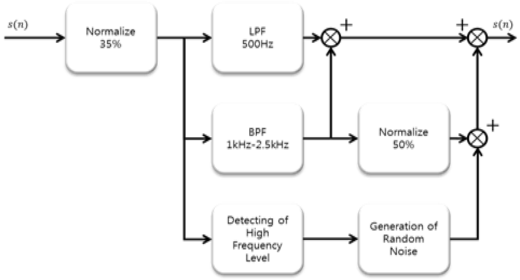


그림 4. 제안된 과부하 음성신호 보상 과정 블록도
Fig. 4. Block diagram of proposed overload speech signal compensation

The first step is to normalize the clipped input signal by about 35%. In order to divide and compensate the high frequency band more precisely than the conventional method, the low frequency, the mid frequency, and the high frequency component are divided and compensated and then synthesized. Pass low-pass filter of about 500Hz to compensate the low frequency part. To compensate for the clipped signal as in the previous method of Figure 2, a low-pass filter is used because the low-frequency waveform must be made of bone and peaks. Even if the distortion rate is large, information for compensating is remained more than the frequency band higher than about 500 Hz. The second step is to use a bandpass filter with a cutoff frequency of about 1000 Hz to 2500 Hz to compensate for the mid-frequency portion. Finally, in order to use the high frequency level of 2500Hz or more of the speech signal, the high-band part is separated through the high-pass filter and compensated by reconstructing it with Gaussian noise. The experimental environment for the proposed method is a 8000 Hz sampling of negative samples and quantization with 16 bits.

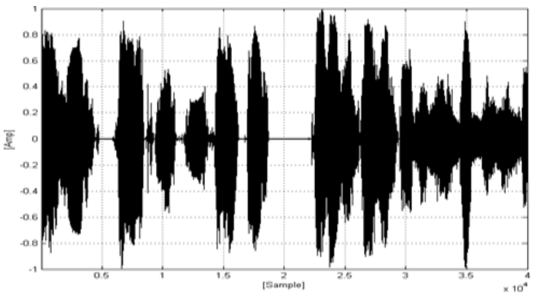


그림 5. 제안된 방법으로 보상된 음주 전 음성파형
Fig. 5. Compensated before drinking speech waveforms by the proposed method

Figure 5 shows the waveforms of Figure 2 before drinking and compensated through the proposed method. From Figure 2, it can be seen that Figure 5 is compensated for in the (+) region as a whole. Figure 6 and Figure 7 show the possibility of drinking by using the method of judging the possibility of drinking through the difference in the LPC degree of each speech signal before and after the compensation.

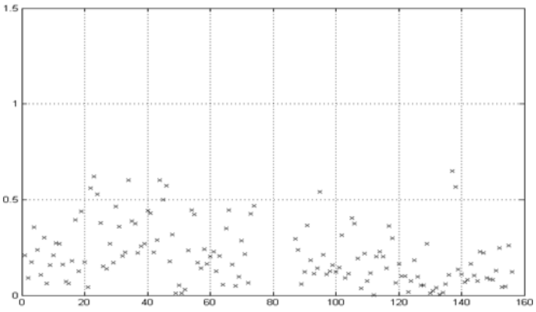


그림 6. 음성 파형 보상 전 음주 가능성 여부 확인을 위한 LPC계수 차이 그래프
Fig. 6. LPC coefficient difference graph for confirming possibility of drinking before speech waveform compensation

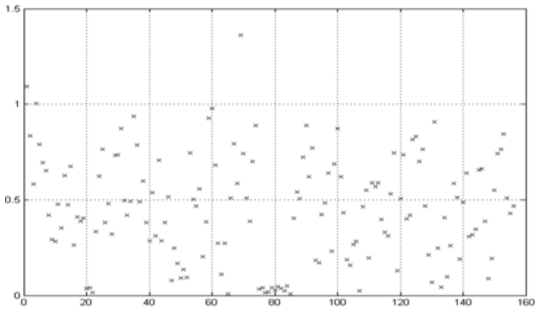


그림 7. 음성 파형 보상 후 음주 가능성 여부 확인을 위한 LPC계수 차이 그래프

Fig. 7. LPC coefficient difference graph for confirming possibility of drinking after speech waveform compensation

When the formants are compared through the speech signal before and after drinking, the changes appear in the third and fourth formants after drinking. Before drinking, the difference between 16th and 4th is obvious when calculating formants using LPC, but there is not much difference after drinking. Therefore, it is possible to determine before and after drinking through the LPC coefficient. Comparing Figure 6 and Figure 7, Figure 7 shows that the difference values according to LPC coefficients are higher than Figure 6. Figure 6 shows the LPC coefficient difference with the pre-compensation speech waveform. The results of Figure 6 show that there is distortion in the 3rd, 4th Formant information, which is a high-frequency region that can detect the possibility of drinking, and it is difficult to judge accurately. Figure 7 shows that the LPC coefficient difference is prominent in each frame as shown in Figure 6 because the information of 3rd, 4th formants is compensated through the proposed algorithm. Can be obtained. Through this, the probability of pre - drinking was increased through speech signal after compensating before compensation. Therefore, it can be seen that the compensated speech signal through the proposed method reduces the error about possibility of drinking.

Ⅳ. Conclusion

Drinking accidents are occurring in the water as much as alcoholic accidents on the road. Drinking accidents occur on the sea, so it is difficult to prevent them beforehand, and if they happen, the cost and damage will increase. If it is possible to find fishing vessels that are operated after drunkenness through remote control, it is possible to enforce intervention efficiently, which can reduce cost and secondary damage. If the possibility of drinking can be judged by using speech data that can transmit a person's body change over a long distance, it is possible to sufficiently intercept the communication only in the Wireless Communication environment. In the case of a cargo ship or a large ship, it is possible to communicate with good sound quality because it uses a specially designed device and has less influence on the noise inside the ship. Conversely, in the case of a small fishing boat or an inexpensive wireless device, a noise or a speech signal may be clipped because the input level is limited. Therefore, in this paper, we proposed a method to compensate the signal to reduce the error rate of the possibility of drunk by distorting the speech signal.

In the future, this information can be used as a criterion for obtaining important information changes in the field of recognition and synthesis, and it can be applied as a measure to reduce errors in the judgment of possibility of new drinking, and as a standard.

References

[1] Statistical Yearbook of National Security Agency, 2016.

[2] Traffic accident statistics for 2016 (National Police Agency), 2016.

[3] 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," International Journal of Engineering and Technology(IJET), Vol.7, No.5, pp. 1957-1964, Oct-Nov 2015.

[4] Seong-Geon Bae, and Myung-Jin Bae, "On A New Hybrid Speech Encoder using Variable LPF," ASTL, Vol. 21, pp. 275-278, April 2013.

[5] Seong-Geon Bae, Myung-Sook Kim, and Myung-Jin Bae, "On Enhancement Signal Using Non-uniform Sampling in Clipped Signals for LTE Smart Phones," in Proc IEEE ICCE-berlin, pp. 125-126 2013.
DOI: https://doi.org/10.1109/ICCE-Berlin.2013.6698030

[6] Seong-Geon Bae, Myung-Jin Bae, "A New Speech Coding using Harmonics Emphasis Filter," in Proc ISAAC 2013, AACL Vol. 1, pp. 43-44, 2013.

[7] Chang-Joong Jung, Seong-Geon Bae, Myung-Sook Kim and Myung-Jin Bae, "Speech Sobriety Test Based on Formant Energy Distribution," IJMUE Vol.8, No.6, pp.209-216 2013.

[8] S. Marcieli Bellé, Sílvia do Amaral Sartori, Angela Garcia Rossi. "Alcoholism: effects on the cochleo-vestibular apparatus". Rev Bras Otorrinolaringol. 73(1):116-22. 2007.
DOI: https://doi.org/10.14257/ijmue.2013.8.6.21

[9] Seong-Geon Bae, and Myung-Jin Bae, "Reducing errors of judgment of intoxication in overloaded speech signal," International Journal of Engineering and Technology(IJET), Vol.8, No.1, pp. 219-223, Feb-Mar 2016.
https://www.scopus.com/inward/record.uri?eid=2-s2.0-84960489412&partnerID=40&md5=1841aa0af36df0afe4b79fd875bef5d7

[10] Seong-Geon Bae, Won-Hee Lee, and Myung-Jin Bae, "A new judgment of intoxication based on changes in low frequency characteristics of speech signal," International Journal of Engineering and Technology(IJET), Vol.8, No.5, pp. 2242-2245, Oct-Nov 2016.
DOI: https://doi.org/10.21817/ijet/2016/v8i5/160805011

[11] Seong-Geon Bae, and Myung-Jin Bae, " On a new hybrid speech coder using variables LPF," International Journal of Software Engineering and its Applications(IJSEA), Vol.7, No.6, pp. 391-398, 2013.
DOI: https://doi.org/10.14257/ijseia.2013.7.5.13

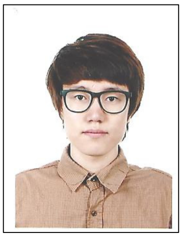
[12] Seong-Geon Bae, and Myung-Jin Bae, " A study on a new non-uniform speech coding using the components of separated by harmonics and formants frequencies," I Communications in Computer and Information Science (CCIS), Vol.352, pp. 246-251, 2012.
DOI: https://doi.org/10.1007/978-3-642-35603-2_36

[13] H.W Park and M.J. Bae, "Study of Annoyance in Relation to Exposure Time to Demonstration Nois," The Journal of The Institute of Internet, Broadcasting and Communication, Vol.16, No.6, pp.103-108.
DOI: https://doi.org/10.7236/JIIBC.2016.16.6.103

[13] Park, H.-W., Bae, S.-G., Bae, M.-J., "A study on drinking judgment by the pitch analysis of speech signal," (2016) International Journal of Engineering and Technology, 7 (6), pp. 2304-2308.
https://www.scopus.com/inward/record.uri?eid=2-s2.0-84954429411&partnerID=40&md5=9ec4d79c7fbf4e7b72b3127eb0999027

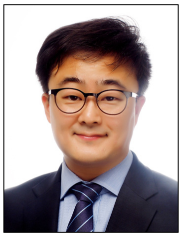
저자 소개

이 원 희(준회원)



• 2012 년 8월 ~ 2014년 8월 : 숭실대학교 정보통신공학과 석사
• 2015 년 3월 ~ 현재 : 숭실대학교 정보통신공학과 박사과정
<주관심분야 : Speech Signal Processing, Acoustic Signal Processing, pattern recognition, Deep Learning>

박 형 우(정회원)



• 2015년 ~ 현재 : 숭실대학교 전자정보공학부 교수
• 2013년 ~ 2015년 : 한국철도기술연구원 선임연구원
• 2013년 : 숭실대학교 정보통신공학과 (박사)
<주관심분야 : 음성 및 음향 신호처리, 소음저감 시스템>

배 성 근(정회원)

- 2016년 ~ 현재 : 강남대학교 소프트웨어응용학부 교수
- 2012년 ~ 2016년 : 대림대학교 방송음향영상학과 교수
- 2014년 : 숭실대학교 정보통신공학과 (박사)

<주관심분야 : 음성 및 음향 신호처리>

배 명 진(정회원)

- 1992년 ~ 현재 : 숭실대학교 전자정보공학부 교수
- 1986년 ~ 1992년 : 호서대학교 전자공학과 교수
- 1987년 : 서울대학교 전자공학과 (박사)

<주관심분야 : 음성 및 음향 신호처리, 소리공학>