International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 11 (2019), pp. 2045-2047

A Study on High Quality Personalized Analysis Method using Improved Speech Bandwidth and Weighted Formants

Seong-Geon Bae[1] and Myung-Jin Bae[2,*]

[1]School of Software Application, Kangnam University, Gyunggido, Korea.
E-mail: sgbae@kangnam.ac.kr, ORCID ID: 0000-0003-3252-0062

[2]Information and Telecommunication of Department, Soongsil University, Seoul, Korea.
E-mail: mjbae@ssu.ac.kr, ORCID ID: 0000-0002-7585-0400

Abstract

This paper is to classify the user's voice bandwidth more precisely so that it can be applied to various application cases. Recently, it is very useful to provide high quality voice service with recognition rate when using artificial intelligence or big data. Especially, considering the characteristics of high sound quality, the bandwidth of sound is adopted at the existing high frequency of 16 kHz, so that the use of data increases, but it is a great advantage to ensure that the high-quality bandwidth can be obtained. In general, the translation width feature applies only necessary data, so it cannot be used in high-quality service. Therefore, this study is to provide personal bandwidth for bandwidth to provide various services using high quality bandwidth.

Keywords: High-quality bandwidth, High frequency, LPC, Pitch, Formant

1. INTRODUCTION

Recently, the demand for artificial intelligence speakers has increased and various services through voice signals have been provided, and for this reason, the demand for high-quality voice in various fields has increased. In the field of voice service, it is divided into analysis and recognition, and in the field of analysis, efforts are made to apply it to analysis and stone through diversification of voice information. Until now, most of the voice information was used and applied using limited bandwidth using formant and basic frequency through the method analysed in the 19th century. The speech analysis method through this analysis does not overcome the limitation of bandwidth in the demand of the times. Since bandwidth limitation has the validity of data, it becomes difficult to provide high quality service if the previous bandwidth is used. Therefore, this study aims to further improve the understanding of high-quality service for information by expanding bandwidth required for high-quality in existing limited band and applying weights to the most important voice band. The problem is that it is difficult to express the user authentication or the specificity of the sound quality when using it as a limit of the band as a feature of the limited bandwidth coming from the existing method, so it is difficult to apply various applications due to the limit of the amount of information. In order to emphasize high quality data and individuality in voice signals, it is necessary to emphasize clear emphasis or characteristics in the pre-treatment process. The study of applying high-quality data is very urgent in the times, which is essential for the problem of strengthening the security of individuality or recognition of speciality [1][4][5][10][11][12][13]. It is important to strengthen various means used for authentication through information to meet the demand of high quality and to obtain information using existing format and basic frequency analysis. Therefore, if the individual's unique voice characteristics are used as improved bandwidth through speech analysis and used for statistical analysis through DB, it can be more accurate and excellent application. This method improves the existing bandwidth of 4kHz and uses weighted weighting method. Considering the increase of data amount, it is applied to the comparison using 5kHz to 8kHz. In this paper, we have improved the bandwidth for voice analysis that can be used for high-quality service, and we have applied the format weight for improving the sound quality. In particular, LPC analysis was used to obtain the formal components in the general use of the basic parameter extraction process, and the analysis method was applied to minimize the influence of the basic frequency affecting the sound quality weight [1][2][3][4][5][6][7].

The proposed method is composed of DBs using the characteristics on the frequency spectrum with the individual characteristics and the feature analysis is used. In chapter II, the basic frequency and formant analysis, which are generally used in voice signals, are described, and in chapter III, the study method on the improved sound width is explained, and in chapter IV, the conclusion and future research direction are suggested.

[*] ***Corresponding author:*** Myung-Jin Bae, Professor.
Department of Information & Telecommunication,
Soongsil University, 369 Sangdo-Ro,
Dongjak-Gu, Seoul, Korea.
E-mail: mjbae@ssu.ac.kr

2. CHARACTERISTICS OF EXISTING AND PROPOSED METHODS

The basic method for speech analysis is possible by analyzing the source and applying it as a filter. This method has been performed using basic frequency change or measurement of the formant frequency. Until now, the parameters used in speech recognition or speech analysis have been used separately by the individual's own frequency component, pitch, and formant, a common resonance frequency component of sound. This is because it is limited to existing bandwidth, and it is difficult to use high-quality service in the gate. Basic frequency with individual characteristics is used as important data for speaker recognition and plays an important role in distinguishing individual characteristics. The method of extracting this is separated and detected in the time domain and the frequency domain. In the application using this parameter, it is excellent as a parameter to recognize and synthesize sound as information about personal identification, but it is not appropriate to identify the clarity or color of sound. Formant information is unusual in the important part of vowels in pronunciation and is characterized by the common sound rather than the individual [1][2][3][4].

This formant parameter is a characteristic that is prominent in vowels and is analyzed as an important part of the recognition or synthesis of notes. Especially, the characteristics of the formant are more important in high quality, and the maintenance of the fourth or fifth formant is the subject of research. The analysis of the natural frequency and formant was extracted using the existing method, which is to minimize the error through the comparison of the analysis method. The methods studied now include energy extraction, zero-crossing, and autocorrelation methods in time domain, and LPC, FFT, and Cepstrum methods in frequency domain. In this study, we used the frequency domain LPC to extract the pouch and formant.

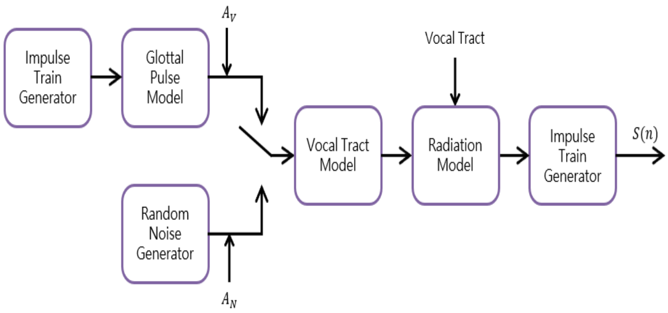
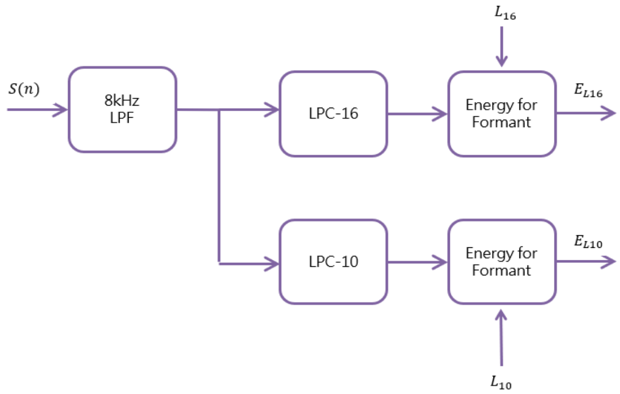


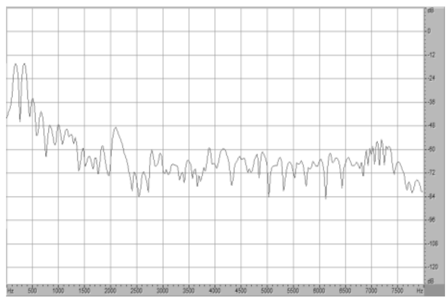
Figure 1. A general parameter extraction method

Basically, to distinguish the characteristics of the sound, the pitch information and the format information are separated and used for synthesis or coding or modeling it. Figure 1 is a picture to explain it. Figure 1 is a picture for explaining the sound generation or generation. The part of the sound generation or generation is designed by using a filter. Since the basic frequency is the best to express the characteristics of the individual, it can be used as an impulse filter used for general pitch generation. Formant modeled the tube and used it for the filter so that the peak part with the highest energy can be the maximum of 4 to 5 depending on the bandwidth.

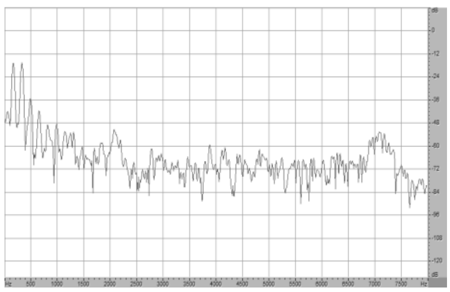
3. PROPOSED METHOD

Generally, to analyse the voice, it is detected by distinguishing the basic frequency and the formant. This study emphasizes the characteristics of Formant and modify it to match the characteristics of high quality, so existing Formant extraction method was used. Especially, as shown in the picture, the order of LPC is detected 10 times according to the existing bandwidth, and the second is the LPC 16th order is extracted to improve the bandwidth. The compared formant was excluded from the improvement because there is no change without the difference in the characteristics of the sound quality. The more the result value of the extracted formant is different, the more the speech improvement is needed, so the improved formant is adjusted. The adjusted bandwidth affects the recognition or change of the sound because most of the order is high.





a) The method for applying the existing method



b) formant weighted value

Figure 2. The proposed method

In this paper, we measured the -3dB part based on the maximum gain of the speech converted to the frequency domain, and we used the bandwidth to obtain the bandwidth and the obtained bandwidth to compare relative. The pitch width obtained in this way can be obtained importantly according to the clarity of the sound, and this information can be found to be important for recognition or identification. The inputted signal is converted into a frequency region and the maximum gain is set to a reference (0dB) to measure and dB the band width at -35 dB and use. The speech signal was adjusted to be used as a parameter at 4kHz~6KHz; the bandwidth expansion at the general 4kHz was used as an improved format weight. Since the existing method is the same as the forced deletion of the individual clarity of the sound, the frequency width of the sound is measured by expanding to 8kHz in the actual study. The experiment and result were input 16-bit data by 16KHz sampling by interface of AD/DA converter for input/output of voice signal. The performance evaluation of the processing results was used for the experiment using the speech signals of teenagers, twenties, thirties, and forties.

5. CONCLUSION AND FUTURE RESEARCH DIRECTION

In this paper, we propose an analytical method that can provide high quality service using improved personalized speech bandwidth, and the results are verified by comparing with existing methods. The basic frequency is not applied to the original sound because of the specificity of the sound. Formant is an important parameter for high-quality service because of its large energy. This is extracted using LPC, which is the basic method, and the characteristics of the formant energy are applied as weights by applying the LPC order differently. The study measured the information of individuality tone more widely than information in limited bandwidth, and thus, to provide high quality service, the characteristics of individuals were used according to the emphasis of Formant and studied as an important factor. In the future, when this information is applied to various applications, it is studied to be available in the existing band limited bandwidth sound quality and to be available in high-quality bandwidth. This study can be used as various parameters in the field of voice recognition and voice synthesis and can be used as a DB applicable to new services such as artificial intelligence

REFERENCES

[1] S.G. Bae, “A Study on Bidirectional Control Parameters using Audio Signals Based on an Adaptation of a Virtual Reality Space for User’s Emotion Change,” International Journal of Applied Engineering Research, Vol. 12/23 (2016), 14796-14801.

[2] S.G. Bae, M.S. Kim, and M.J. Bae, “On Enhancement Signal Using Non-uniform Sampling in Clipped Signals for LTE Smart Phones,” 2013, IEEE ICCE-berlin, pp.125-126, ICCE-berlin 2013.

[3] Won-Hee Lee, Myung-Sook Kim, and Myung-Jin Bae, Using valid-frame deviation to judgment of intoxication. Information: An International Interdisciplinary Journal, Vol. 18 (2015), 4131-4136.

[4] Seong-Geon Bae, Myung-Jin Bae, “A New Speech Coding using Harmonics Emphasis Filter,” ISAAC 2013, AACL Vol. 1(2013), pp43-44.

[5] Myungjin Bae, Myungsuk Kim, Professor Bae’s Sound Story, Gimm-Young Publishers, Inc, 2013.

[6] 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, pp.235-237.

[7] Seong-Geon Bae, Won-Hee Lee and Myung-Jin Bae, " A Study on Low Frequency Noise of Dehumidifier using Acoustic Charactristics," IJET Vol.8, No.1, pp.235-237.

[8] Won-Hee Lee, Myung-Jin Bae, “Reducing Errors of Judgment of Intoxication in Overloaded Speech Signal,” IJET Vol.8, No.1(2016), pp.219-224.

[9] Seonggeon Bae, Myungjin Bae, “A New Speech Coding using Harmonics Emphasis Filter,” ISAAC 2013, AACL Vol. 1(2013), pp43-44.

[10] Yeong Kyoung Su, Bae Myoung Jin, "Psychoacoustic study of automobile engine sound", The Acoustical Society of Korea, Vol.34, No.2, pp.180, 2015.

[11] 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.

[12] 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.

[13] Zwicker, E., Fastl, H. (1990): Psychoacoustics - Facts and Models, Springer - Verlag, Berlin.