Study on quality improvement of bone-conductive voice

Tomohiro Minami, Keisuke Nishida, Takahito Ujike and Shigeaki Aoki*

Department of Electric, Information and Communication Engineering, Kanazawa Institute of Technology,
7–1 Ohgigaoka, Nonoichi, 921–8501 Japan

(Received 9 April 2012, Accepted for publication 31 May 2012)

Keywords: Bone-conductive voice, Voice quality, Background noise

PACS number: 43.60.Dh [doi:10.1250/ast.33.376]

1. Introduction

As one of the methods for sending a voice in telecommunications under a high-noise environment, a bone-conductive microphone is effective. Since the bone-conductive microphone records voice vibrations transmitted in the skull, its anti-noise characteristic is superior to that of air-conductive sound transmitted in the air. However, bone-conductive sound has the drawbacks of reduced sound quality and articulation score. It is considered that these drawbacks are mainly caused by the loss of high-frequency components. Therefore, it is expected that the quality of bone-conductive sound is improvable by appropriately generating these high-frequency components.

In this paper, a method of restoring the lost high-frequency components of bone-conductive sound by utilizing an alias spectrum is proposed. Moreover, we conduct a listening test to confirm the effectiveness of the proposed method.

2. Proposed method

2.1. Application of an alias spectrum

As a method of reproducing the missing high-frequency region from an original signal with a restricted frequency region, the spectrum extrapolation method using the alias spectrum generated by resampling has been proposed [1]. This method is effective for generating the frequency region that is absent in the original signal. It does not need a prerecorded database of both voices for the design of filters and is able to manipulate parameters in actual use. In the case of bone-conductive sound, the high-frequency components are decreased in intensity. Therefore, it is expected that the method can be applied to the improvement of bone-conductive sound. We investigate the possibility of this application.

2.2. Analysis of bone-conductive sound

In applying the spectrum extrapolation method using the alias spectrum generated by resampling to bone-conductive sound, it is necessary to decide the frequency range in which the bone-conductive sound is effective. To clarify this, using a bone-conductive microphone (G-450) and an air-conductive microphone (NL-31), both the bone-conductive and air-conductive sounds of an adult male’s voice are recorded simultaneously by a PC with a USB audio interface (Sound BLASTER Digital Music SX). The bone-conductive micro-

*p = e-mail: aoki.s@neptune.kanazawa-it.ac.jp

phone is inserted in the external auditory canal and the air-conductive microphone is set 15 cm in front of the mouth. The sampling frequency is 44.1 kHz. Then, their frequency components are analyzed.

A typical example of the frequency spectra obtained is shown in Fig. 1. It shows that the level of the bone-conductive sound in the high-frequency region is decreased clearly in comparison with that of the air-conductive sound. The effective frequency region of the bone-conductive sound was not defined typically as a low-pass filter. However, in order to clarify the influence of the boundary frequency, it was generally judged that the effective frequency was below about 2 to 2.5 kHz.

2.3. Configuration of system

From the analysis in Sect. 2.2, which showed that the effective frequency region of bone-conductive sound is limited to below about 2 to 2.5 kHz, three different system configurations were considered for use with the spectrum extrapolation method. In these three configurations, the boundary frequency between the original signal and the alias spectrum, that is, the folding frequency, is set at 2, 2.5 and 3 kHz.

Block diagrams of the three system configurations are shown in Fig. 2. The sampling frequency \( f_s \) of the input signal is 8 kHz. The alias spectrum is generated by performing suitable combinations of upsampling and downsampling processing. The alias spectrum becomes the frequency components in the high-frequency region. Eventually the generated signal is resampled with 8 kHz sampling frequency. The anti-aliasing filter processing is included in the operation of \( P/Q \) (\( P, Q \): arbitrary integers) in the block diagram.

The schematic frequency characteristics of the bone-conductive sound generated using the proposed method are shown in Fig. 3. Using the boundary frequencies of 2, 2.5 and 3 kHz, it is shown that the frequency characteristics of the high frequency region differ and the degree of emphasis of the high-frequency region can be adjusted.

3. Results and discussion

3.1. Test method

Subjective tests using the method of paired comparisons are conducted on a total of four different signals [2]. They are an original bone-conductive voice and three different processed voices whose boundary frequencies are 2, 2.5 and 3 kHz. The voices used for the tests are recordings of three

---

*PACS number: 43.60.Dh [doi:10.1250/ast.33.376]

---

**Keywords:** Bone-conductive voice, Voice quality, Background noise
texts read by two adult males that are recorded using the bone-conductive microphone. The three texts are from phonetically balanced speech data distributed by NTT-AT. A pilot test was conducted including an air-conductive voice. The win rate of the air-conductive voice was so high that it was excluded from the subjective test.

The combinations of pairs that can be chosen from four voices is six, and if permutations are taken into consideration, the number of permutations becomes 12. Each combination is presented five times. Therefore a total of 60 random combinations are presented to three subjects. The subjects answer by the question, “which sound has better quality?” In addition to the above-mentioned processing, all test signals are passed through a band pass filter. The bandwidth is 0.2–3.8 kHz.

3.2. Test results and consideration

The results of paired comparisons are shown in Table 1, where A is the original bone-conductive voice and B, C and D are the processed voices whose boundary frequencies are 2, 2.5 and 3 kHz, respectively. In the results, consistency was satisfied and the agreement was significant ($\alpha = 0.05$).

It is clear that the percentages of wins of C and D, which were processed, were much higher than those of A, which was not processed. B had the lowest percentage of wins, which was even lower than that of A. The percentage of wins of C shows that C had the most effective frequency, which was followed by D.

The results showed that the boundary frequencies of 2.5 and 3 kHz of cases C and D, respectively, were effective. However, there was a difference between C and D. The alias frequency regions in C acted effectively. On the other hand, because there were slightly fewer alias frequency regions in D, it is thought that the effect of processing was small. Moreover, in the case of B, hardly any improvement in sound quality was found. It is considered that since the alias frequency regions were too wide, the natural properties of the voice were lost.

4. Conclusion

In this paper we investigated the quality improvement of bone-conductive sound by using a proposed method. The method restores the loss of bone-conductive sound by...
utilizing an aliasing spectrum and resampling. The subjective test result showed the validity of sound quality improvement by this proposed method. It turned out that an optimal boundary frequency for processing by aliasing existed, which was 2.5 kHz in the subjective test.

We plan to modify the preprocessing by changing the boundary frequency, adding an equalizer and so forth, and to add postprocessing for the removal of artificiality to further improve the sound quality.

References