A Study On The Analysis of Sound Signals According To Adaptive Sound Image Change In Various Virtual Spaces

Seong-Geon Bae

School of Software Application Kangnam University, Youngin-si, Gyungg,y Korea
sgbae@kangnam.ac.kr

Abstract - *Contents made with virtual spaces are evolving in various ways. Using the features of virtual spaces, they create avatars between each other even from a distance to express their personalities. The characteristics of virtual spaces it is further developed by creating diversity and developing individual characteristics. These days, using the metaverse spaces, are being used as spaces where individuals and a large number of people meet. These spaces give us freedom in terms of time and space that we have to meet face-to-face. The complex and varied considerations of face-to-face can undermine the purpose and delivery of content. Therefore, this study considered the characteristics according to various changes by considering the acoustic characteristics in the metaverse environment with these evolving virtual spaces.*

Keywords - *Virtual spaces, metaverse, acoustic characteristics, adaptive image filter*

I. INTRODUCTION

Various characteristics appearing in virtual space are to consider visual image effects and acoustic characteristics. Visual characteristics should be processed in consideration of the characteristics of image quality or frame rate. This is because these characteristics affect the fatigue of the eyes and the realism felt in the virtual space. Recently, the metaverse using avatars has been applied and further developed. It is developing as a non-face-to-face solution rather than a visual effect. These environmental characteristics are being understood and studied. However, since it is very important to give realism to the characteristics of audio if it is applied in consideration of the characteristics of audio in the surrounding environment, it is possible to approach the realism even more and to eliminate the sense of distance in spatial movement. These characteristics have not been studied yet. Considering the specificity of the image in the virtual space and solving it with picture quality and frame, it should be emphasized more in the audio aspect as it changes in various ways with technological development and research. In order to solve these characteristics, in this study, the feedback method was applied and applied to the environment was used. This method was studied due to the adaptation time required for environmental change. Since the human ear is sensitive to the environment and adapts the balanced characteristics according to the continuity of the surroundings, the time to adapt to the background sounds appearing in the virtual space was used through feedback rather than directly applied to the environment. The method used in this study is a method of minimizing distortion by improving the method of adapting to the acoustic environment and feeding back the method of adapting to the visual change. Chapter 2 introduces the existing methods and understands their features. Chapter 3 explains the theory applied using the proposed method. Chapter 4 shows the experiment and results using the proposed method. Finally, Section 5 describes a conclusion.

II. EXISTING METHOD

In general, most of the sensory effect methods used in virtual space make good use of visual characteristics. In most studies, it is not necessary to apply audio, and the characteristics are not considered because they focus only on dialogue with each other in real space. For this reason, acoustic properties are used only in dialogue in the metaverse used for virtual space or non-face-to-face use. When using in a non-face-to-face special space, the characteristics of the two-way must be carefully considered because the focus is on the conversation without considering the characteristics of the surrounding sound, which are essential for each person's environment. Furthermore, these characteristics should be taken into account because it is considered a muting phenomenon or a network anomaly during the conversation. Considering the environmental sound that gives a sense of space in the movement of space in virtual space or in augmented reality is because it makes good use of the characteristics of realism. Therefore, the existing acoustic considerations simply stayed at the level of sound effects. This is because these sound effects are intended to focus on scene changes or necessary scenes in the metaverse space. Therefore, the phenomena appearing as these problems are that the sense of reality is lowered, concentration is significantly lowered, and can only be achieved in one direction in a conversation. In order to solve these problems, in the previous study, the frame change method and the energy difference change method were applied, and the sound effect and the sound of the real space were additionally added. The frame change method is a method of matching the sound effect to the frame by detecting when the temporal scene suddenly changes with the characteristics of the sound effect. This is a method of giving a realistic effect by selecting various sound effects according to visual changes and mixing them during playback, rather than simply continuing playing

sound effects. Figure 1 shows the frame change method. Since the frame change method is to reproduce various sound effects according to the change of the scene, there is a disadvantage in that the amount of calculation is increased when the frame is changed slowly or the image quality is high. To compensate for these shortcomings, an energy comparison method was used.
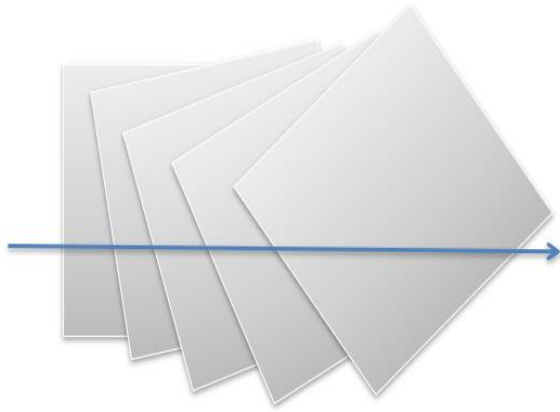
(1- Weighting) x image

Figure 1 Existing frame change method

The energy difference change method is a very effective method to change the sound effect in various ways when the energy difference of the scene is greatly changed among the characteristics of the screen. Figure 2 explains how to change the energy difference. This method is effective for realism and concentration, but in the metaverse space, the flow of dialogue is not considered, and the visual scene is selected as a visual scene, so the effectiveness is poor. In virtual or augmented reality, since the characteristics of various sound effects are expressed as energy, they respond well to sudden changes, but there may be problems with scenes with slow changes or special characteristics. Since most of these existing studies depend on changes in the environment, the concentration in the virtual space or the concentration in the metaverse space can be reduced, but the realism is reduced, and it is difficult to consider the scene in the non-face-to-face, especially in the metaverse space. It is not easy to apply. In addition, since the sense of realism and the sense of space or realism felt in the metaverse space cannot be conveyed, a method to solve these problems is required. In this study, to improve this problem, Lee Ji-mi's adaptive realism enhancement was studied. Each of the existing methods has advantages and disadvantages. The frame comparison method has a huge amount of computation, so the calculation required for real-time application is huge, and it is very useful for abrupt image changes. The histogram corporate energy change method can be applied by reducing the step size when more sophisticated detection is required by probabilistic comparison of the histogram for each frame. These methods have been applied in previous studies and obtained various results.
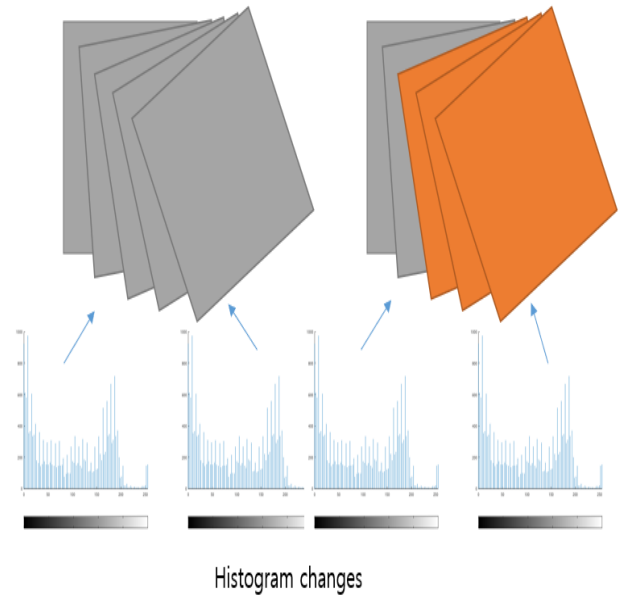

Histogram changes

Figure 2 Conventional energy comparison method

III. PROPOSED METHOD

The most important part of virtual space is image processing. These days, it is giving a sense of realism in a variety of ways due to the original development, but there is a lack of research on the contents considering the acoustic characteristics. Therefore, in this study, in order to solve the problem of matching the image in the frame change method of the acoustic part and the energy comparison method, feedback on the individual environment and ambient sound effects are added to help maintain various immersion and specificity of dialogue. Give.

$$En(u) = \sum_{n=1}^{k} (image(n) - image(n-1))^2$$

(1)

The formula used in this study extracts the characteristics of the image input from each frame by calculating the deviation from the energy or histogram. First of all, the characteristic of the metaverse, which is sensitive to the personal environment, is that the immersion in the dialogue takes into account the characteristics of noise and sound in different spaces. Even in a virtual space, the maladaptive environment sound cannot be created because the user's environment is taken into consideration. Second, in the method of mixing the ambient sound about the environment with the virtual space sound, the feedback was selected by giving a certain amount of time. This was used in consideration of the specificity that the ears accept.
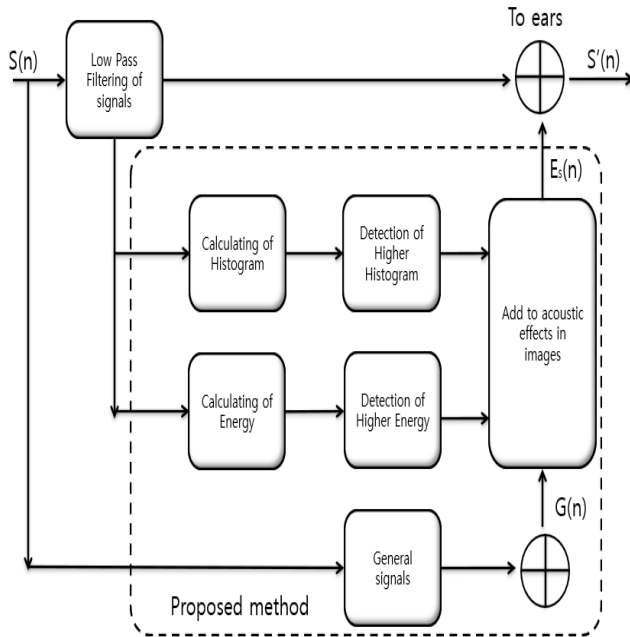
Fig. 3. A Newly proposed method

In the figure, the conventional frame change metric and the energy change metric, a histogram, are used. These methods were introduced in previous studies and detect sharp changes in images. These characteristics were used by applying the existing method. The frame change method can detect features that appear in the abrupt image change section well, and the energy change method detects the change section in 30 frames using statistical values from the histogram. Since this detected method has advantages and disadvantages, two methods were used with variations. Here, a feedback method was applied to minimize the influence on the surrounding environment. In this case, the user's anxiety about the environment or each other's environment is synthesized, so it is considered to exist in reality. The proposed method is applied and used in various virtual spaces and metaverse spaces, and since the environmental noises of the virtual space and real space, which are each other's environments, are properly synthesized, it can be used naturally in augmented reality. Even in today's metaverse space, the sense of reality facing each other grows as if they are facing each other by synthesizing the sounds of each other's environments, which were shared only in virtual space.

ACKNOWLEDGEMENT

IV. CONCLUSION
This study is to develop an image-adaptive realistic sound that can be used in virtual space or metaverse space. In general, it was difficult to develop a sound that adapts to various environments or image changes because sound effects are used regardless of changes to the image or focus only on dialogue. We overcame the shortcomings of law and energy change and conducted applicable research. Because this study emphasizes realism more, it makes it possible to better adapt to the user's adaptation and the surrounding environment because acoustic considerations are applied as feedback in the metaverse space or virtual space. In the future, these studies can be well adapted to virtual space or augmented reality and can be applied to stable environmental changes because the surrounding environment of real space is additionally applied. The additional feedback environment can have the effect of reducing dizziness and heterogeneity by reducing the adaptation time even when the user returns to the real space. In the future, it is considered that more effective research will be made if a data DB is created and a more precise sound for sound effects is added in spite of such environmental changes.

REFERENCES

[1] Seong-Geon Bae., A Study on the Improvement of Sound Quality by Compensatory Variables through Improved Openness in Virtual Space, IJETT, 68(11) (2020) 89-91.

[2] Seong-Geon Bae., A Study on the Acoustic Variation to Environment Changes in Virtual and Complex Reality IJERT, 13(4) (2020) 648-651.

[3] Seong-Geon Bae and Myung-Jin Bae., A Study on Recovery in Voice Analysis through Vocal Changes before and After Speech Using Speech Signal Processing, IJAER, 12 (2017) 5299-5303 .

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

[5] Seong-Geon Bae and Myung-Jin Bae., A Study on Recovery in Voice Analysis through Vocal Changes before and After Speech Using Speech Signal Processing, IJAER, 12 (2017) 5299-5303.

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

[7] M. J. Bae and S. H. Lee., Digital Voice Signal Analysis, Books Publishing Dong Young, (1998) 3,6.

[8] Seong-Geon Bae, A Study on the Acoustic Variation to Environment Changes in Virtual and Complex Reality, IJERT, 13(4) (2020) 64-651.

[9] Seong-Geon Bae, Won-Hee Lee, and Myung-Jin Bae., A Judgment of Intoxication using Hybrid Analysis with Pitch Coutour Compare in Speech Signal Processing, IJAER, 12 (2017) 2342-2346.

[10] Seong-Geon Bae and Myung-Jin Bae., A Study on Recovery in Voice Analysis through Vocal Changes before and After Speech Using Speech Signal Processing, IJAER, 12 (2017) 5299-5303.

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

[12] Seong-Geon Bae, Myung-Sook Kim, and Myung-Jin Bae, Using High-Frequency Accentuation in Speech Signals as a New Parameter in Intoxication Judgment. Information, An International Interdisciplinary Journal,17 (2014) 6531-6536.

[13] Seong-Geon Bae, Myung-Jin Bae, A New Speech Coding using Harmonics Emphasis Filter, ISAAC 2013, AACL 1 (2013) 43-44.

[14] Geum-Ran Baek, Myung-Jin Bae, Study on the Judgment of Intoxication State using Speech Analysis . 352 CCIS, (2012) 277-282, Springer-Verlag.