Design and Implementation of a Voice Feedback Device for Voice Loudness Control

Fumiya Hara, Yoshinari Takegawa, Keiji Hirata

Design and Implementation of a Voice Feedback Device for Voice Loudness Control

Fumiya Hara[1⋆], Yoshinari Takegawa[1], and Keiji Hirata[1]

[1]Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan
{g2117040,yoshi,hirata}@fun.ac.jp
http://www.hiratakelab.jp

Abstract. In recent years, wearable devices such as Google Glass and Apple Watch are spreading in society. However, their devices cannot be used the following types of situations. In the case of user wants to quieten children being noisy in a public place, the user wants to speak clearly in a presentation despite being nervous. Therefore, wearable devices are important for the user to control his/her actions. This research aims to achieve involuntary and non-perceptual control of the user's actions, without inflicting mental stress. The first stage of the research is to design and implement a voice feedback system for voice control. This device is composed of an environmental sound measurement microphone, a speaking voice input microphone, a switch for the white noise function, a minicomputer, and a volume amplifier circuit. A first function of the device change the user's voice feedback from earphones by the user's voice loudness. Moreover, a second function of the device that output white noise add in order to consciously increase the use's voice loudness. In this paper, I describe the proposed vocal volume control and the implementation method of the proposed system.

Keywords: Loudness Feedback System · White Noise System · Wearable Device · Involuntarily

1 Introduction

In recent years, information provision techniques that consider the user's situation are attracting notice in the field of wearable computing. Google Glass and Apple Watch are such wearable devices that provide beneficial information for the user. For example, navigation to a destination, and health management based on number of steps and heart rate. These information provision techniques present information for choices of action and promote specific actions. However, there are situations in which we cannot immediately control our own actions, even when it is vital. For example, in the case of user wants to quieten children being noisy in a public place, user wants to speak clearly in a presentation despite being nervous, a disaster or other unexpected event. Therefore, wearable devices are important for the user to control his/her actions. This research aims

⋆ Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan

to achieve involuntary and non-perceptual control of the user's actions, without inflicting mental stress. The first stage of research is to design and implement a voice feedback system for voice control.

This research applies the effect by which humans involuntarily speak in a loud voice in a noisy environment [1]. Moreover, changes the voice loudness that the user hears (hereafter: Heard voice loudness) according to the environment sound and controls the voice loudness at which the user speaks (hereafter: Speaking voice loudness).

2 Relevant Research

Since many years ago, a great amount of research on human vocalization models has been conducted. Speech Chain [2] is a major example of this. Humans control voice loudness according to auditory information, such as ambient sound, and visual information, such as the distance between oneself and the listener. Also, Lombard Effect is a voice model specialized in auditory feedback. We naturally tend to use loud voices in noisy environments, and listening to this noise not only causes the voice to become loud but is also known to change a variety of acoustic features, for example by causing the fundamental and formant frequency to rise. There are various research results relating to Lombard Effect [1]. For example, Hodoshima, Arai, Kurisu [4] surveys clarity of voice in a silent environment, a noisy environment, and a reverberant environment. Our research differs from existing research in that it surveys the effects of providing user voice feedback in real time. On the other hand, there are many existing research cases on control of user's action by wearable devices.

Various approaches to involuntary control of action exist, including control of appetite by visual presentation by VR (Virtual Reality) [5], and control of action by vibrating motor[6, 7].

Kurihara, Tsukada [8] developed a system called SpeechJammer which inhibits speech, without imposing physical pain, by using Delayed Auditory Feedback to artificially delay when the speaker hears his own speech. SpeechJammer is closely related to our research in terms of control of speech. The approach of this research is different in that in aims for involuntary and non-perceptual control of user's speech.

3 Designs

3.1 Principle of Control of Voice Loudness

It is thought that the user's voice loudness can be changed by controlling the information the user can obtain from sound and sight. One of the reasons that voice control depends on auditory information is thought to be the relationship between environmental sound and one's own voice loudness. For example, our voices are louder than usual when we talk in a noisy place such as a concert hall or construction site. This phenomenon is called Lombard Effect [1]. Conversely, our voices are quieter than usual when we are in a quiet place such as library. In other words, we compare voice loudness with environmental noise and appropriately control voice loudness to maintain the difference between the loudnesses of the two sounds. In the case of noisy places, a person compares environmental noise and their ordinary voice loudness, and finds the voice loudness to be low, as a result of which voice loudness increases. In the case of quiet places, voice loudness becomes low because it is found to be louder than environmental sound.

Following on, it is thought that perspective is involved as a visual approach. Our voice loudness increases when we want to talk to a person who is far away. In contrast, our voice loudness decreases when we want to talk to a person who is nearby. Using this, this research designs and implements a basic voice loudness control system and system structure.

3.2 Voice Loudness Control System

To propose our voice loudness control system, we built the hypothesis that speaking voice volume becomes quiet when heard voice volume is loud, and, inversely, speaking voice volume becomes loud when heard voice volume is quiet. The reason for this is described below.

Amplification of Voice Loudness To amplify the user's voice loudness, first, it was considered that the current heard voice volume is made quieter than the normal heard voice loudness. Second, the user will increase speaking voice loudness in order to make the level of the current heard voice loudness the same as the normal heard voice loudness. Therefore, we must make the current volume of environmental noise louder than the usual volume of environmental noise that the user hears with their own ears, in order to attenuate heard voice loudness. Accordingly, we considered that, in the situation where heard voice loudness is amplified, attenuating heard voice loudness will result in amplification of speaking voice loudness.

Attenuation of Voice Loudness To attenuate the user's voice loudness, we must invert voice loudness amplification. That is to say, we considered that the current heard voice loudness is amplified the current speaking voice volume will be involuntary attenuated.

3.3 Usage Scenarios

In the Case of Increasing Voice Loudness There are few people who do not feel stress during interviews for job hunting and part-time jobs, examinations. In particular, stress gets worse as the importance of the interview increases. By preventing this problem and controlling the voice loudness transmitted to the user, it is possible to eliminate the issues of what the interviewer is saying being inaudible or not understood. Also, by inducing appropriate voice loudness in the same manner as in the interview scenario, a presenter can use this system to efficaciously make a presentation to listeners.

In the Case of Decreasing Voice Loudness It is poor manners to speak loudly when other people are nearby, such as during a party at someone's house, in a library or on an aeroplane. However, we sometimes forget to be considerate when having an exciting conversation. It is expected that our system, which causes the user to control voice volume involuntarily, can reduce voice loudness, even when the user has forgotten to can considerate and without altering the user's consciousness.

3.4 Design Policy

Involuntariness and Non-perception There are some situations in which we consciously change voice loudness. For example, when a presenter speaking quietly while making a small-scale presentation is told to speak more loudly, or when someone talking loudly in a library is warned to become quieter. After receiving such a warning, we attempt to change voice loudness, but tend gradually to return to the former voice loudness so that there is no change. The system proposed by this research provides the user with heard voice loudness, as mentioned in section 3.1, in order to resolve these problems.

This system needs to be a designed in such a way that the user does not notice changes to heard voice loudness, so that they may speak without perception. Thereby, the user is given heard voice loudness matched to environmental sound. Also, when the degree of amplification changes it is not changed suddenly, but at such a speed that the user does not notice.

Response to Continually Changing Situation and Environment Environments and situations in which the proposed system can be used include quiet environments, such as a library, and noisy environments, such as a party. The system must constantly acquire environmental sound and return voice feed-back appropriate for the environmental sound at each moment. Therefore, it is necessary to consider the validity of the stages of amplification, in order to give appropriate heard voice loudness.

Also, when giving a presentation, the system feeds back white noise when the user consciously wishes to increase speaking voice loudness. This is because, by the Lombard Effect [1], the user's speaking voice loudness increases when the loudness of white noise increases. Accordingly, it is thought that the system can be utilized in situations such as public speaking. As a result, this research implements a switcher that changes between a feed-back function and a white noise function.

Miniaturization and Weight Reduction The wearable device created in this research is supposed to be used in as many different situations and environments as possible, thus miniaturization and weight reduction are essential. Therefore, this wearable device only focuses on heard voice loudness and does not use visual recognition, enabling the system to be constructed from only microphone, earphones, and a minicomputer such as Arduino.

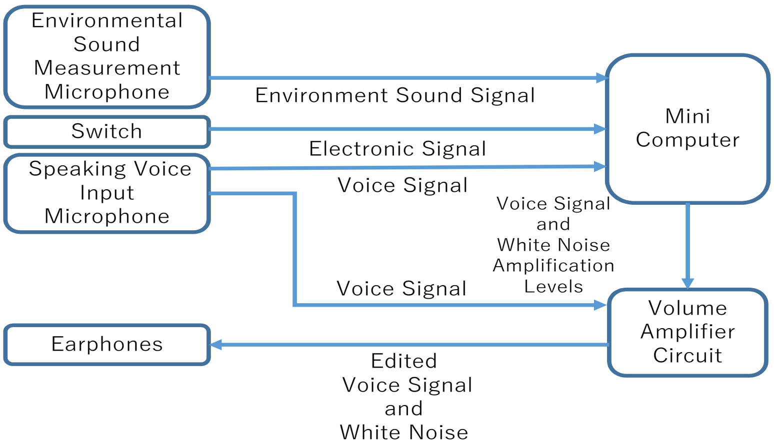

Fig. 1. System configuration

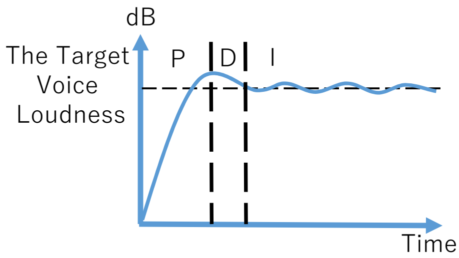

Fig. 2. For example of amplification of voice loudness

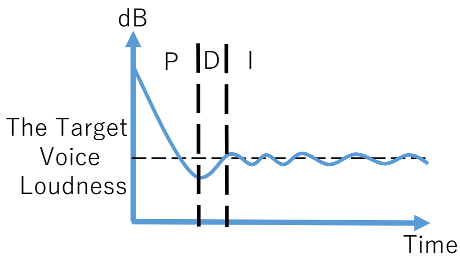

Fig. 3. For example of attenuation of voice loudness

Table 1. Amplification of voice volume feed-Back

Levels	Pulse	Volume Amplification
1	550 μ s	− 12dB
2	1050 μ s	− 6dB
3	1500 μ s	0dB(basic point)
4	1950 μ s	+6dB
5	2400 μ s	+12dB

Table 2. Amplification of white noise

Levels	Amplification
1	0dB(basic point)
2	+2dB
3	+8dB
4	+12dB
5	+16dB

3.5 Implementation

The system [1] configuration of our wearable device is shown in fig 1. This device is composed of an environmental sound measurement microphone, a speaking

[1] You can watch a demonstration video at https://1drv.ms/f/s!AqUbzynYF8M8ijk0JO0Rnx3p4AAp

voice input microphone, a switch, earphones, a minicomputer, and a volume amplifier circuit. First, the system decides the amplification level of the volume amplifier circuit based on the environmental sound, the user's voice loudness and the target voice loudness programmed in advance using the minicomputer.

To establish target voice volume, we established five ranks of amplification of loudness. These ranks are shown in table 1. PID (Proportional-Integral-Differential) control [9] is used as a control model of an amplification of loudness depending on environment sound. Moreover, the reason for establishing five ranks for amplification of loudness is to inspect how much speaking voice loudness changes in relation to heard voice loudness. First, -12dB is the state in which voice feed-back is the smallest, and is close to the level of an ordinary conversation. However, as explained in section 3.2, speaking voice loudness increases when heard voice loudness decreases, and speaking voice loudness decreases when heard voice loudness increases. To prove that speaking voice becomes quieter, we first give heard voice loudness level 3 when using our device. The system switches to amplification level 2 when the user wants to increase speaking voice loudness in a loud environment by decreasing heard voice loudness. If the user wants to speak even more loudly, the amplification level switches to 1. The control model of an amplification voice loudness is shown in fig 2. First, the user is given the heard voice loudness for the user's voice loudness to reach the target voice loudness. An amplification voice loudness is equivalent to a proportional control(P). Moreover, an integral control(D) is to perform voice loudness control that does not exceed the target voice loudness. Finally, a differential control(I) is to perform voice volume control so as to eliminate the difference from the target voice loudness. Next, amplification switches to level 4 when the user wants to decrease speaking voice loudness in a quiet environment. If the user wants to speak even more quietly, the amplification level switches to 5. The control model of an attenuation voice loudness is shown in fig 3. A proportional control(P) with relation to an attenuation voice loudness is shows the opposite effect to an amplification voice loudness.

Also, we implemented an additional function that plays white noise to make speech louder when users want to increase voice loudness by themselves. This function sends a signal from the microcomputer and increases white noise when a switch is pressed. At this time, voice loudness feed-back is not output. We have established five levels of white noise amplification, and the level increases each time the switch is pressed. The amplification levels are shown in the table 2 below. The control model is as same as figure 1 and the heard voice loudness takes the place of white noise. When the user is using this functions but does not speak, white noise is not output, so as not to impede the speech of others. We will design a wearable system as described in this paper, and conduct an evaluation experiment.

References

1. Lane,H. and Tranel,B. The Lombard Sign and the Role of Hearing in Speech, J. Speech Hear. Res, Vol. 14, pp.677-709, 1971.

2. Denes,P. and Pinson,E. The Speech Chain, 2nd Ed, New York : W.H Freeman and Co. 1993.
3. S, Hiroya. T. Mochida. M, Kashino.: Articulatory gestures, not auditory frequency resolution, determine formant frequency discrimination thresholds in vowels, Abstracts of The 29th MidWinter Meeting of Assoc. Res. Otolaryngol, 2006.
4. Hodoshima, N. Arai, T. Kurisu, K: Intelligibility of speech spoken in noise and reverberation , Proceedings of 20th International Congress on Acoustics (ICA2010), pp. 3632-3635, 2010.
5. Narumi, T. Ban, Y. Kajinami, T. Tanikawa, T. Hirose, M.: Augmented perception of satiety: controlling food consumption by changing apparent size of food with augmented reality, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 109-118, 2012.
6. Rekimoto, J.: Traxion: A Tactile Interaction Devicewith Virtual Force Sensation, Proceeding of Annual ACM Symposium on User Interface Software and Technology (UIST2013), pp. 427-432. 2013.
7. Junji, W. , Ando, H. Maeda, T. : Shoe-shaped Interface for Inducing a Walking Cycle,Proceedings of the 2005 international conference on Augmented tele-existence (ICAT2005), pp. 30-34, 2005.
8. Kurihara, K. Tsukada, K: SpeechJammer: A System Utilizing Artificial Speech Disturbance with Delayed Auditory Feedback, The 18th Workshop on Interactive Systems and Software, pp.77-82, 2010.
9. Paul Avery:IntroToPIDControl.pdf, Introduction to PID control(online), available from (http://www.machinedesign.com/sites/machinedesign.com/files/datasheets/gated/IntroToPIDControl.pdf) (accessed 2017-06-19)