IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 March 2023, accepted 9 April 2023, date of publication 21 April 2023, date of current version 1 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3269067

METHODS

Remote-Control Sound Pattern Recognition for Auditory Assessment and Therapy in Smart Healthcare

TAK CHUN KWONG[1], **(Student Member, IEEE), YAT SZE CHOY**[1], **AND STEVE W. Y. MUNG**[2], **(Senior Member, IEEE)**
[1]Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong
[2]Research and Development Office, The Education University of Hong Kong, Hong Kong

Corresponding author: Steve W. Y. Mung (wymung@eduhk.hk)

This work was supported in part by the Innovation and Technology Commission under Project UIM381, and in part by the Start-Up Research Grant from The Education University of Hong Kong under Project RG 29/2022-2023R.

ABSTRACT A remote-control method is proposed for healthcare services involving auditory assessment and therapy through video-conferencing software. Different video-conferencing software platforms have different sound responses, and the user computer also has variations in sound performance because of its hardware and software configuration. These restrictions affect the ability of video-conferencing software to deliver health services that require accurate, reliable, and repeatable sound. In this paper, a remote-control system board is developed wherein sound for auditory assessment or therapy is stored. A low-cost processor is used to remotely deliver sound by sound pattern recognition. Service providers play a sound pattern locally to remotely control the system board. The desired audio files can be delivered to the client side without suffering from sound loss in terms of frequency range and loudness. Experiments were also conducted to verify the performance of the proposed method. The results show that the remote-control method for audio playback can be realized with a low-cost processor, such as a microcontroller unit for sound pattern recognition and that it accurately delivers audio files in terms of frequency and amplitude.

INDEX TERMS Healthcare, microcontroller unit, pattern recognition, telehealth.

I. INTRODUCTION

The global Coronavirus (COVID-19) pandemic has greatly impacted our society and different measures have been implemented to ensure the human safety. Social distancing is a measure that has been commonly adopted during the pandemic. As a consequence, healthcare services with face-to-face settings have been forced to search for a new means of delivery to ensure client and provider safety. Telehealth is a smart healthcare option [1], [2], [3] that is able to cope with this challenge via well-developed internet and corresponding tools. Telehealth provides healthcare services from professionals using information and communication technologies when distance is a critical factor. These services advance the health of individuals and their communities through the exchange of valid information for diagnosis, treatment, and prevention of diseases and injuries, continuing education of healthcare providers, and research and evaluation. Telehealth can be employed in a synchronous manner, such as via real-time video-conferencing software, or in a store-and-forward way, wherein the exchange of information is performed asynchronously. These two approaches can also be used simultaneously to provide relevant services [4], [5]. The utilization of telehealth models is beneficial to people who have difficulty traveling or live in remote areas as it is a viable alternative to in-person services. This creates the possibility for clients to access expert services unavailable in nearby regions and enables the provision of services despite a long distance. In addition, telehealth could potentially save travel time and delivery costs for clients and service providers [6], [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.

Telehealth has been used in various fields prior to the COVID-19 pandemic. Music therapy is an example of a therapeutic service that employs telehealth [8], [9], [10]. Music therapy is the clinical use of music intervention to achieve individualized goals within a therapeutic relationship by a credentialed music therapist [11]. In the past 20 years, effects of sound and music therapy on patients [12] has become research of interest. Music therapy can be used to address physiological, psychological, and spiritual needs of patients [13]. Previous studies have suggested that music therapy has an impact on anxiety [14], [15], sleep disorders [16], [17], depression [18], [19], [20], stress reduction [21], [22], [23], [24], and psychiatric diseases [25], [26]. It has also been revealed as an alternative to pharmacological methods to relieve pain and disturbance due to its non-invasive nature and lack of side effects [27], [28], [29]. Music-based therapy corresponds to two fundamental methods: an "active" method based on playing musical instruments and a "receptive" method based on listening to music [30]. The frequency range of the music used in therapy covers the audio spectrum from 20 to 20000 Hz.

Apart from therapeutic service, health assessment services like audiometry have also integrated telehealth into their service delivery [31], [32], [33]. Basic audiometric tests involve the use of pure tone and speech stimuli to perform routine screening and diagnostic procedures [34]. For example, pure tone audiometry determines the lowest sound intensity at each frequency that a client can hear [35]. The result reveals the hearing threshold of the client, and this method could be used to identify possible hearing deficits [36]. Due to the global prevalence of hearing loss, there exists a large group of people who have difficulty accessing hearing healthcare services because of various barriers. Therefore, the application of telehealth in audiometry offers a potential solution through use of information and communication technologies [37]. Under a pandemic, there is an even higher demand to remotely deliver these services so that patients and clients can continue to receive interventions or assessments without bearing additional risk.

With the advancement in technology and rising number of smartphone owners [38], web-based and app-based remote hearing assessment tools have become more accessible [39], [40]. The combination of portable and smart devices such as tablets and laptops with remote hearing assessment tools offers opportunities for accessible hearing healthcare services to populations worldwide. In addition, in a recent review [41], 187 remote hearing assessment tools in the form of smartphone apps and online platforms were identified in the market. However, their reliability and accuracy varied drastically across tool categories, with one of the issues being the need for calibration on different devices.

There are several challenges in realizing the ideal generation of sound over video-conferencing software. In telehealth services involving video-conferencing software to deliver sound, it is crucial for the system to accurately generate the original sound since sound intensity is animportant parameter in some services. The first challenge is that video-conferencing software limits the frequency spectrum of the audio signal. In a preliminary report written by Howell et al. [42], four evaluated platforms showed that the software attenuates high frequency sound at the receiver end at a cutoff frequency range from 7000 to 12500 Hz. Additionally, alteration of sound intensity and the introduction of inharmonic noise is uncontrollable due to the audio configuration of client computers and headsets.

In this paper, a remote-control method is proposed to enable service providers to remotely play desired audio files at the client side without loss of sound frequency range and loudness due to the conferencing platform and the client computer. A remote-control system board was developed for the proposed method. This control method is suitable for providing health services involving auditory assessment and therapy through video-conferencing software. The contributions and background information corresponding to the work described in this paper are discussed in Section II. The methodology of the proposed remote-control method is described in Section III and its implementation is presented in Section IV. Section V discusses and evaluates the experimental results of the system board in terms of delay, sound amplitude, and sound frequency. Finally, Section VI explains potential study contributions.

II. BACKGROUND AND CONTRIBUTION

Variation in sound output level of different electronic devices is one problem that hinders the reliability of telehealth services, such as hearing tests. Although studies have shown that similar brand smartphones have no significant difference in output level [43], [44], there is a maximum difference of 8 dB in audio output level across six smartphone devices from four major smartphone manufacturers [45]. For a personal computer, output-level variation also exists due to differences in available audio hardware. The measurement setup of audio output amplitude response versus audio input amplitude of audio MP3 files played by a computer is shown in Figure 1. A commercial laptop was used for this measurement and the MP3 files were sinusoidal signals generated at octave frequencies in the range of 250 to 8000 Hz with amplitude ranging from 0.1 to 10 mV with 0.1-mV step size. The audio output was connected to an audio analyzer, and, in Figure 2, the results show that the audio output of the laptop is not proportional to the audio input and there is saturation at the medium level at some frequencies. Although audio files can be calibrated and measured beforehand by an audio analyzer, there is no guarantee that each computer will deliver the desired audio response. Client headsets are driven by these electrical signals and the speakers in the headset convert this electrical signal to sound waves. The amplitude of the electrical signal determines the sound pressure level (dBSPL). The hearing level (dBHL) is the sound intensity commonly used in audiology and this dBHL comes from the sound pressure

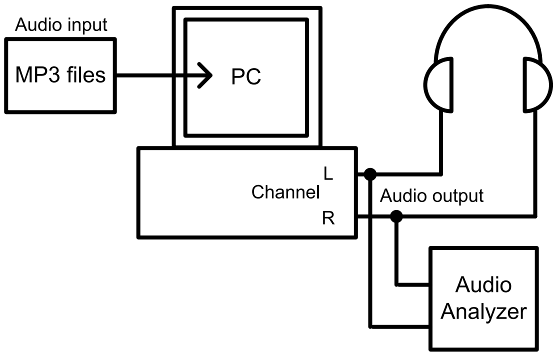


FIGURE 1. Measurement setup for comparing audio output amplitude from a computer versus audio input amplitude from MP3 files.

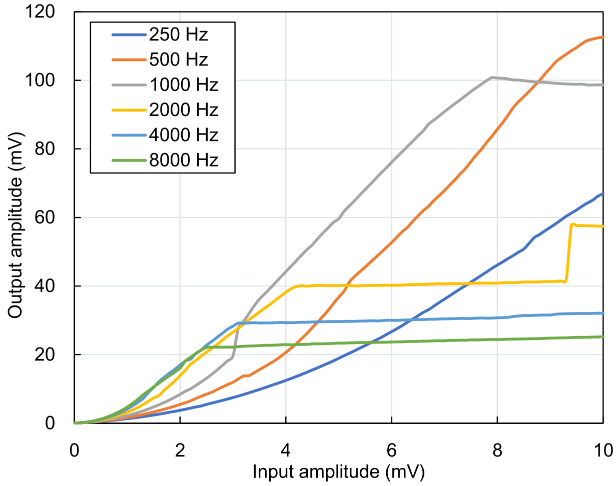


FIGURE 2. Audio output amplitude of a commercial computer versus input audio file amplitude.

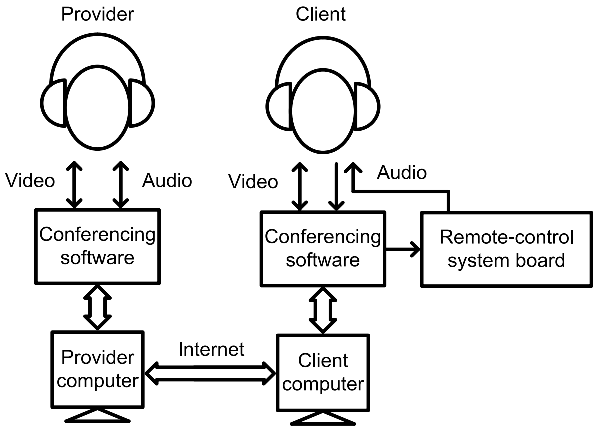


FIGURE 3. Remote setup of the proposed remote-control method.

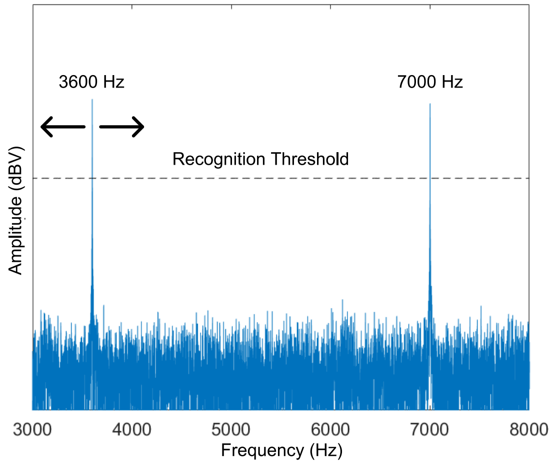


FIGURE 4. Frequency spectrum of a frequency pattern example composed of tonal frequencies at 3600 Hz and 7000 Hz.

level (dBSPL) generated from the speakers, i.e.,

$$\mathit{dBHL} = \mathit{dBSPL} - \mathit{thresholds} \tag{1}$$

The reference level for dBHL is zero, which is related to the threshold of dBSPL for average people and differs at different frequencies. Therefore, the difference between the hardware and operating system used in different computers will lead to sound intensity deviations from the electrical signals, even when the same headsets or speakers are used.

This paper proposes a remote-control method with the following characteristics:

- Service providers can use the method to control desired audio files played at the client side without any loss of sound frequency range and loudness because of hardware and software used in the service.
- A remote-control system board is proposed to implement the proposed method by sound pattern recognition.
- A low-cost processor, such as a microcontroller unit (MCU), is proposed to implement sound pattern recognition.

III. METHODOLOGY

The remote setup for the proposed method is shown in Figure 3. At the provider side, there is no alteration of the video-conferencing software usage. The audio and video paths were identical to the normal settings. At the client side, a change is made at the audio path by inserting the remote-control system board between the computer and the headset. The headset is calibrated together with the system board as a playback system using head and torso simulators to ensure optimized audio characteristics and frequency response. This system board includes a processor, audio codec, and memory for storage. It is responsible for remote switching of the audio signal between conferencing and playing audio test files by the service provider. This operation is realized by performing recognition of the incoming audio signal from the client-side computer. The audio files are saved in the memory and each file is assigned to a unique frequency pattern, which can be speech, music, or other multitones.

A simple unique frequency pattern is used in the example shown in Figure 4. This basic frequency pattern is composed of two tonal frequencies. The higher frequency, 7000 Hz, is constant for all patterns and acts as the trigger frequency. The lower frequency is the selective frequency, referring to a particular audio file in the memory that is selected by changing the selective frequency to a particular frequency below the trigger frequency. A fast Fourier transform (FFT) operation is performed continuously on the audio signal coming from the video-conferencing software. After each operation, the magnitude of the frequencies used to form the patterns are also compared to a recognition threshold (Figure 4). If both magnitudes are greater than the threshold, the playback function is triggered and the audio file corresponding to that frequency pattern is played. If it is not recognized, the pass-through function keeps running for communication. Since the health service provider and client will still be using the video-conferencing software for communication, the threshold magnitude is set to a higher level to avoid false activation of the playback function. To control the playback remotely through this conferencing platform, the provider sends the sound signal with the frequency pattern assigned to the corresponding audio file through the video-conferencing software. When the system recognizes the unique sound signal sent from the provider side, it will switch from conferencing audio to the playback of the audio test file. The simple tonal frequencies shown in Figure 4 can be used as the frequency pattern. As each audio file has its unique pattern, remote control of playback of individual files can be accomplished. Other complicated patterns, such as speech, could be used for recognition so that more audio files could be played remotely and the possibility of false triggering playback by speech and environmental noise presented in the input signal can be further minimized.

Since one of the objectives of the proposed method is to provide a smooth transition between audio from the video-conferencing software and playback audio when a particular frequency pattern is detected, it is essential that minimal delay is introduced in the process. There are two types of delays in operation: pass-through delay, which is additional delay during normal use of video-conferencing software due to the processor, and playback delay, which is the amount of time required to switch from the conferencing software audio signal to the playback of audio files. An experiment was designed to evaluate the pass-through delay resulting from the proposed method. Sine bursts were generated by the audio analyzer at octave frequencies in the range of 250 to 16000 Hz with a magnitude of 100 mV. The signals were passed to the audio input of the system board, and both signals at the audio input and output of the system board were captured by the audio analyzer. The audio signals were recorded three times and then averaged to calculate the pass-through delay.

A second experiment was designed to evaluate the feasibility of the proposed method and to measure the playback delay. A set of frequency patterns with a trigger frequency set as 7000 Hz and a selective frequency ranging from 1000 to 6800 Hz with a 200-Hz step was used. A computer from the service provider was used to generate a constant 7000-Hz background sinusoidal signal, which was identical for each frequency pattern. Sine bursts were generated in the range of 1000 to 6800 Hz with a 200-Hz step. All signals were generated with a magnitude of 100 mV_{rms} and each pattern was passed to the audio input of the device to activate audio file playback. The signals at the audio input and output of the system board were captured by the audio analyzer. The audio signals were recorded three times and then averaged to calculate the playback delay.

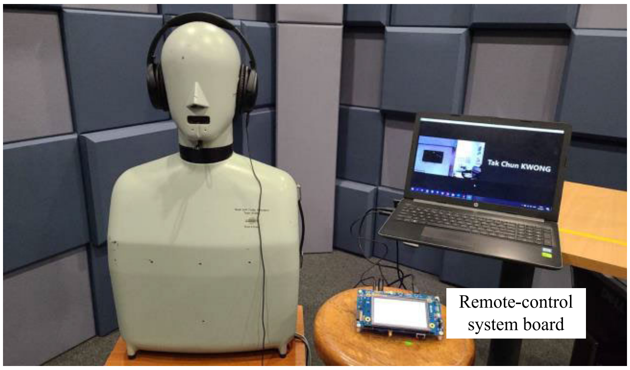


FIGURE 5. Setup of the proposed method at the client side using the proposed remote-control system board.

Apart from the additional delay, it is also important for the proposed method to provide a linear frequency response over the audible range during the pass-through function because most conferencing platforms suffer from limited audio frequency response (as shown in Figure 2). A third experiment was designed to evaluate the frequency response of the device. Frequency sweeps from 20 to 20000 Hz were generated at 1 V_{rms} as the input signal to the system board. The audio input and output of the system board were then captured by the audio analyzer to obtain the frequency response.

IV. IMPLEMENTATION

A commercial board was used as the remote-control system board during verification of the proposed method. This commercial board includes a high-performance mainstream MCU (STM32H735IG) as a low-cost processor that features a double-precision floating point unit and supports digital signal processing instructions. This kit also includes a stereo audio codec with serial audio interface (SAI), two audio jacks for input and output, and a microSD card slot. These features make it suitable for remote control with video-conferencing software since it offers storage and playback of audio files from the memory card and is capable of efficient FFT of incoming signals without disturbing the pass-through function. Figure 5 shows the setup of the proposed method at the client side. Head and torso simulators and ear simulators were used to measure the acoustic performance of *dBHL* from (1).

A block diagram illustrating implementation of the proposed method is shown in Figure 6. The kit is connected to a video-conferencing device to receive audio signal for sound pattern recognition. The operation of the processor in the kit can be divided into three modules: pass-through, FFT, and playback. The pass-through module enables the audio signal to go straight from its input to the output via Path A when no frequency pattern is recognized so that the client can hear sound from the video-conferencing software as usual. During

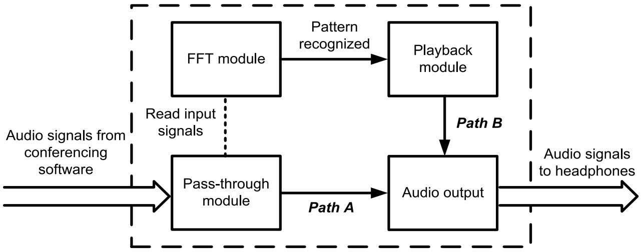


FIGURE 6. Block diagram of the implementation of proposed method in remote-control system board.

this process, the FFT module constantly monitors the input signal and performs operations to detect any designated patterns. When a pattern is recognized, the processor interrupts and disables the incoming audio and activates the playback module to play the corresponding file that is saved locally through Path B. With the proposed method, audio-signal-limiting issues from the conferencing platform can be eliminated. Furthermore, the processor operates independently from the conferencing application. Therefore, no additional software installation is required for the client computer. Once the processor is powered and makes the audio connection, it can work with any device that supports video-conferencing applications, such as a laptop or a smartphone.

Figure 7 shows the program flowchart when performing remote control with the video-conferencing software. After initiating all necessary settings of the direct memory access (DMA) in the general processor and the codec, the DMA is prompted to data communication between the processor and the codec in the SAI frame. This indicates initiation of the pass-through function. The FFT function is executed when there is no callback while continuously checking the starting condition of the playback function. The existence of the target frequency pattern in a mapping table is checked. For example, the simple pattern in Figure 4 is used and has a trigger frequency of 7000 Hz and a selective frequency chosen by the user. The choice of the selective frequency determines which audio file is to be played. When a target frequency pattern is recognized, the DMA and codec are immediately de-initiated to avoid the pattern being heard by the client. The array index of the corresponding pattern is passed to the playback function such that the audio file with the index is played. Finally, de-initiation of the DMA and codec is performed once again before going back to the pass-through function, resulting in completion of the remote-control operation.

During verification, one SAI frame consisted of 64 bits that could be divided into left and right channels so that each channel contains 32 bits. Both channels carry sample data with a 16-bit data size. Therefore, there is a 16-bit space separating each sample data on the bus. A timing diagram of the pass-through module is shown in Figure 8, where A1, A2, A3, B1, B2, and B3 represent the sample data with a size of half buffer. When the full callback fires, three actions occur in parallel. The first action is the continuous update of sample data A2 in the first half of the DMA-received buffer with A2 received from the codec input through the path mentioned in Figure 6. At the same time, the MCU copies B1 in the second half of the received buffer to the second half of the DMA-transmitted buffer. Once it finishes the copying operation, the MCU switches back to perform FFT to check for any designated frequency pattern. The third action is continuous play of A1 in the first half of the DMA-transmitted buffer through codec output. When the received buffer is half-filled with A2, the half callback fires so that the MCU begins to copy A2 into the first half of the DMA-transmitted buffer. The second half of the received buffer is filled with B2 captured from the codec input and the codec plays B1 in the second half of the transmitted buffer. Alternating half and full callbacks complete the operation cycle.

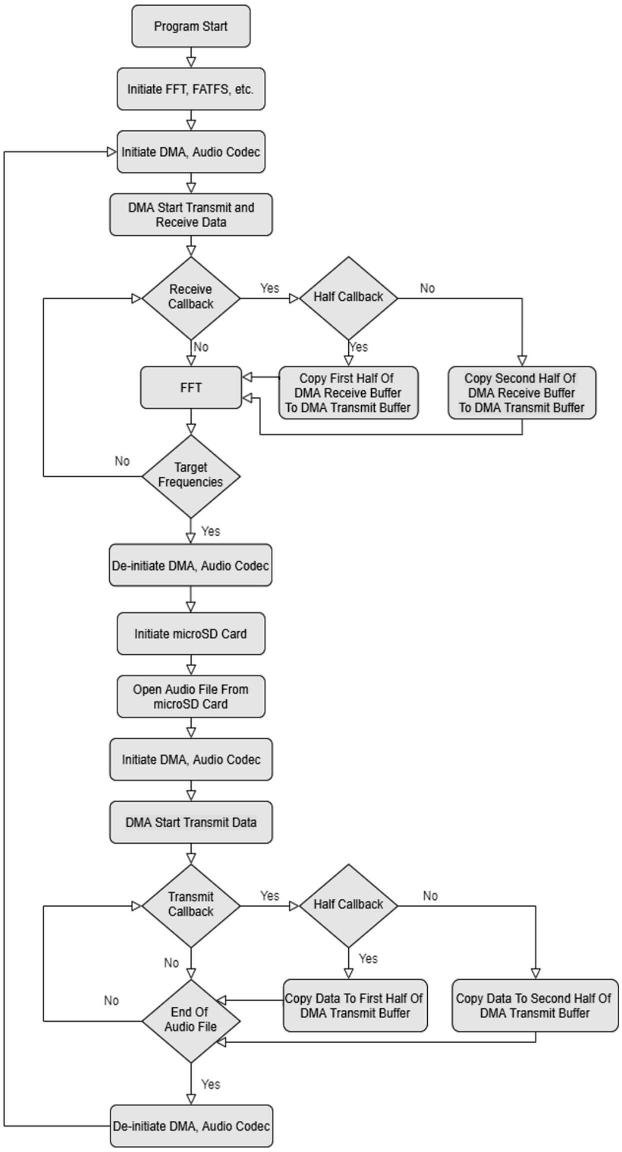


FIGURE 7. Program flowchart of the proposed method.

V. EXPERIMENTAL RESULTS

Table 1 shows the average pass-through delay from 250 to 16000 Hz from the first experiment. The delay was around

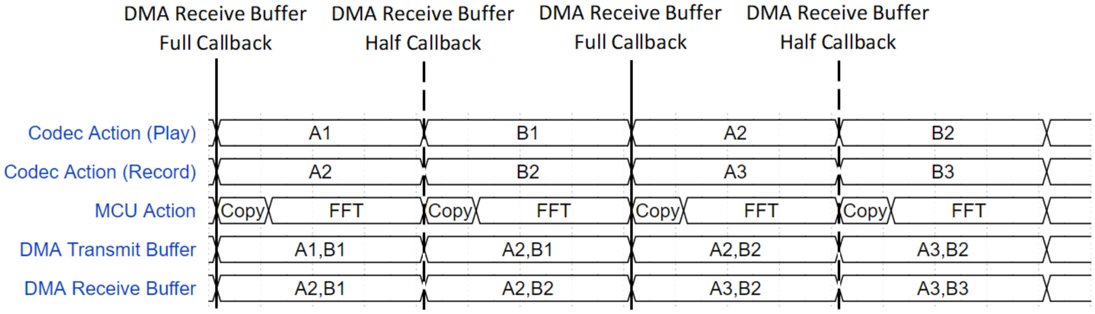


FIGURE 8. Timing diagram of the pass-through module.

TABLE 1. Average pass-through delay of the processor.

Frequency (Hz)	Average pass-through delay (ms)
250	42.31
500	42.31
1000	42.33
2000	42.31
4000	42.31
8000	42.31
16000	42.33

42 ms across the measured frequency. The observed steady delay is due to the method used for handling incoming data. When half of the received buffer is filled, the data is copied to the transmitted buffer and is ready to be played. The delay is the same, regardless of the frequency of the input signal, since the buffer size is fixed. In the synchronization of audio and video, a study has shown that a delay of 80 ms is not noticeable by the majority of subjects and that most subjects responded that annoyance related to the delay is acceptable [46]. Therefore, the additional delay introduced using the proposed method is not significant and will not have any effect on the perceived quality of the conferencing service.

In the second experiment, the MCU successfully recognized all patterns and played corresponding audio files. The average playback delay is shown in Figure 9. The delay obtained across the measured frequency range stays between 564 and 575 ms, which is the normal playback time of the general MP3 player. A higher performance processor can be used to reduce this time. Finally, the frequency response of the low-cost processor is shown in Figure 10. Compared with the input signal, loss of signal amplitude is observed from 20 to 100 Hz. The maximum difference is only around 0.5 dBV, which is relatively small and less likely to cause any noticeable changes in the perceived audio signal. The frequency response stays flat across the 100 to 20000 Hz range.

Experimental results showed that the proposed remote-control method can be implemented to ensure reproduction of audio playback at the client side with control at the service provider side using audio signals with a designated frequency pattern. The reliability of the system is evaluated in the second experiment in which the system successfully switched to audio playback when it recognized the unique frequency pattern and retuned to pass-through of audio from video-conferencing platform after playback. The low-cost MCU implementation also allows this method to be employed without creating a large financial burden for users. Therefore, this verifies the ability of the proposed method to deliver a linear audio signal across the audible range while overcoming

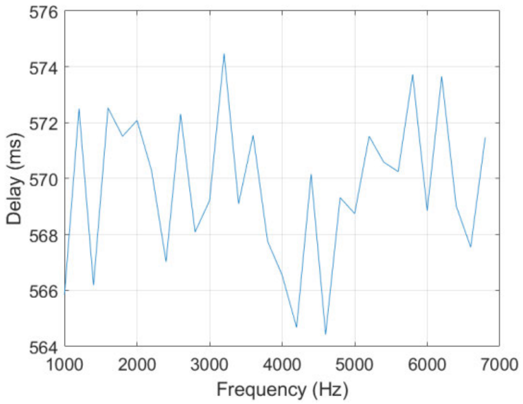


FIGURE 9. Timing diagram of the pass-through module.

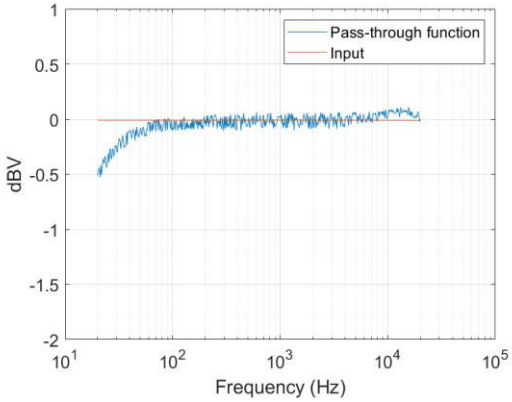


FIGURE 10. Processor frequency response.

limitations observed in most conferencing platforms without relying on the audio files from the internet.

VI. CONCLUSION

In this article, a remote-control method that uses a low-cost processor is proposed for providing health services involving sound through video-conferencing software. The method was evaluated using a remote-control system board in three experiments. The results demonstrate that the output audio signal has a flat frequency response across the audible range. This method ensures generation of accurate, reliable, and repeatable audio playback in terms of frequency and amplitude. Therefore, this method is suitable for smart healthcare applications, such as audiometry, auditory behavioral tests, and music therapy. More complicated sound patterns, such as speech, can be used for recognition to deliver more audio signals in future.

ACKNOWLEDGMENT

(Tak Chun Kwong, Yat Sze Choy, and Steve W. Y. Mung are co-first authors.)

REFERENCES

[1] P. Maji, H. K. Mondal, A. P. Roy, S. Poddar, and S. P. Mohanty, "IKardo: An intelligent ECG device for automatic critical beat identification for smart healthcare," *IEEE Trans. Consum. Electron.*, vol. 67, no. 4, pp. 235–243, Nov. 2021.

[2] P. Chanak and I. Banerjee, "Congestion free routing mechanism for IoT-enabled wireless sensor networks for smart healthcare applications," *IEEE Trans. Consum. Electron.*, vol. 66, no. 3, pp. 223–232, Aug. 2020.

[3] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "ESeiz: An edge-device for accurate seizure detection for smart healthcare," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 379–387, Aug. 2019.

[4] M. Krumm, "Audiology telemedicine," *J. Telemedicine Telecare*, vol. 13, no. 5, pp. 224–229, Jul. 2007.

[5] R. H. Eikelboom and D. W. Swanepoel, "Remote diagnostic hearing assessment," in *Telepractice in Audiology*, E. Rushbrooke and K. T. Houston, Eds. San Diego, CA, USA: Plural, 2016, pp. 123–139.

[6] K. M. Blaiser, D. Behl, C. Callow-Heusser, and K. R. White, "Measuring costs and outcomes of tele-intervention when serving families of children who are deaf/hard-of-hearing," *Int. J. Telerehabilitation*, vol. 5, no. 2, pp. 3–10, Dec. 2013.

[7] P. A. Jennett, L. A. Hall, D. Hailey, A. Ohinmaa, C. Anderson, R. Thomas, B. Young, D. Lorenzetti, and R. E. Scott, "The socio-economic impact of telehealth: A systematic review," *J. Telemedicine Telecare*, vol. 9, no. 6, pp. 311–320, Dec. 2003.

[8] R. Vaudreuil, D. Langston, W. Magee, D. Betts, S. Kass, and C. Levy, "Implementing music therapy through telehealth: Considerations for military populations," *Disabil. Rehabilitation. Assist. Technol.*, vol. 17, no. 2, pp. 201–210, 2020.

[9] H. Spooner, J. B. Lee, D. G. Langston, J. Sonke, K. J. Myers, and C. E. Levy, "Using distance technology to deliver the creative arts therapies to veterans: Case studies in art, dance/movement and music therapy," *Arts Psychotherapy*, vol. 62, pp. 12–18, Feb. 2019.

[10] F. Baker and R. E. Krout, "Therapeutic songwriting with clients in an e-health environment," in *Music Technology in Therapeutic and Health Settings*, W. Magee, Ed. Philadelphia, PA, USA: Jessica Kingsley, 2014, pp. 299–310.

[11] American Music Therapy Association (AMTA). *Definition and Quotes About Music Therapy | American Music Therapy Association (AMTA)*. Accessed: Jun. 13, 2021. [Online]. Available: https://www.musictherapy.org/about/quotes/?print=y

[12] J. Bradt and C. Dileo, *Medical Music Therapy: A Meta-Analysis & Agenda for Future Research*. Cherry Hill, NJ, USA: Jeffrey Books, Cop, 2005.

[13] S. E. Hosseini and S. A. Hosseini, "Therapeutic effects of music: A review," *Rep. Health Care*, vol. 4, no. 4, pp. 1–13, 2018.

[14] J. Bradt, C. Dileo, and M. Shim, "Music interventions for preoperative anxiety," *Cochrane Database Syst. Rev.*, vol. 2013, no. 6, Jun. 2013, Art. no. CD006908.

[15] J. Hole, M. Hirsch, E. Ball, and C. Meads, "Music as an aid for postoperative recovery in adults: A systematic review and meta-analysis," *Lancet*, vol. 386, no. 10004, pp. 1659–1671, 2015.

[16] M. J. Cordi, S. Ackermann, and B. Rasch, "Effects of relaxing music on healthy sleep," *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, Jun. 2019.

[17] L. Andrews, "Music for insomnia in adults," *Clin. Nurse Spec.*, vol. 30, no. 4, pp. 198–199, Jul. 2016.

[18] S. Aalbers, C. Gold, X. Wang, and M. Crawford, "Music therapy for depression," *Cochrane Database Syst. Rev.*, vol. 2017, no. 11, Jan. 2017, Art. no. CD004517.

[19] A. Raglio, "Effects of music and music therapy on mood in neurological patients," *World J. Psychiatry*, vol. 5, no. 1, p. 68, 2015.

[20] M. F. Chan, Z. Y. Wong, H. Onishi, and N. V. Thayala, "Effects of music on depression in older people: A randomised controlled trial," *J. Clin. Nursing*, vol. 21, nos. 5–6, pp. 776–783, Mar. 2012.

[21] A. Linnemann, M. B. Kappert, S. Fischer, J. M. Doerr, J. Strahler, and U. M. Nater, "The effects of music listening on pain and stress in the daily life of patients with fibromyalgia syndrome," *Frontiers Hum. Neurosci.*, vol. 9, p. 434, Jul. 2015.

[22] A. Raglio, D. Bellandi, M. Gianotti, E. Zanacchi, M. Gnesi, M. C. Monti, C. Montomoli, F. Vico, C. Imbriani, I. Giorgi, and M. Imbriani, "Daily music listening to reduce work-related stress: A randomized controlled pilot trial," *J. Public Health*, vol. 42, pp. 81–87, Apr. 2019.

[23] A. Linnemann, B. Ditzen, J. Strahler, J. M. Doerr, and U. M. Nater, "Music listening as a means of stress reduction in daily life," *Psychoneuroendocrinology*, vol. 60, pp. 82–90, Oct. 2015.

[24] F. Giordano, E. Scarlata, M. Baroni, E. Gentile, F. Puntillo, N. Brienza, and L. Gesualdo, "Receptive music therapy to reduce stress and improve wellbeing in Italian clinical staff involved in COVID-19 pandemic: A preliminary study," *Arts Psychotherapy*, vol. 70, Sep. 2020, Art. no. 101688.

[25] M. Geretsegger, K. Mossler, L. Bieleninik, X. Chen, T. Heldal, and C. Gold, "Music therapy for people with schizophrenia and schizophrenia-like disorders," *Cochrane Database Syst. Rev.*, vol. 2017, no. 5, Dec. 2017, Art. no. CD004025.

[26] C. Gold, H. P. Solli, V. Kruger, and S. A. Lie, "Dose–response relationship in music therapy for people with serious mental disorders: Systematic review and meta-analysis," *Clin. Psychol. Rev.*, vol. 29, no. 3, pp. 193–207, Apr. 2009.

[27] U. Nilsson, N. Rawal, B. Enqvist, and M. Unosson, "Analgesia following music and therapeutic suggestions in the PACU in ambulatory surgery; A randomized controlled trial," *Acta Anaesthesiologica Scandinavica*, vol. 47, no. 3, pp. 278–283, Mar. 2003.

[28] A. YousefinejadOstadkelayeh, A. Madadi, S. R. Majedzadeh, R. Shabannia, N. Sadeghian, A. R. Zarinara, S. Sadeghian, and A. R. Jaddian, "The effect of music therapy on chronic pain in patients with cancer," *J. Inflamm. Dis.*, vol. 9, no. 1, pp. 39–42, 2005.

[29] M. M. Y. Tse, M. F. Chan, and I. F. F. Benzie, "The effect of music therapy on postoperative pain, heart rate, systolic blood pressure and analgesic use following nasal surgery," *J. Pain Palliative Care Pharmacotherapy*, vol. 19, no. 3, pp. 21–29, Jan. 2005.

[30] S. Guétin, F. Portet, M. C. Picot, C. Pommié, M. Messaoudi, L. Djabelkir, A. L. Olsen, M. M. Cano, E. Lecourt, and J. Touchon, "Effect of music therapy on anxiety and depression in patients with Alzheimer's type dementia: Randomised, controlled study," *Dementia Geriatric Cognit. Disorders*, vol. 28, no. 1, pp. 36–46, 2009.

[31] D. W. Swanepoel, D. Koekemoer, and J. Clark, "Intercontinental hearing assessment—A study in tele-audiology," *J. Telemedicine Telecare*, vol. 16, no. 5, pp. 248–252, Jul. 2010.

[32] D. W. Swanepoel and L. Biagio, "Validity of diagnostic computer-based air and forehead bone conduction audiometry," *J. Occupational Environ. Hygiene*, vol. 8, no. 4, pp. 210–214, Jan. 2011.

[33] S. Govender and M. Mars, "The use of telehealth services to facilitate audiological management for children: A scoping review and content analysis," *J. Telemedicine Telecare*, vol. 23, no. 3, pp. 392–401, Apr. 2017.

[34] R. J. Roeser and J. L. Clark, "Behavioral and physiological measures of hearing: Principles and interpretation," in *Auditory Disorders in School Children: The Law, Identification, Remediation*, R. J. Roeser and M. P. Downs, Eds., 4th ed. New York, NY, USA: Thieme Medical, 2004, pp. 27–69.

[35] R. W. Harrell, "Pure tone evaluation," in *Handbook of Clinical Audiology*, J. Katz, Ed. 5th ed. Philadelphia, PA, USA: Lippincott Williams Wilkins, 2002, pp. 71–87.

[36] G. J. Forzley, "Audiometry," in *Pfenninger Fowler's Procedures for Primary Care*, J. L. Pfenninger and G. C. Fowler, Eds. 2nd ed. St. Louis, Missouri: Mosby, 2003, pp. 409–415.

[37] C. Brennan-Jones, R. Eikelboom, and D. Swanepoel, "Telehealth for diagnosis of hearing loss," Division Otolaryngol., Univ. Cape Town, Cape Town, South Africa, Tech. Rep., 2021.

[38] *Ericsson*. Accessed: Mar. 23, 2023. [Online]. Available: https://www.ericsson.com/4ad7e9/assets/local/reports-papers/mobility-report/documents/2021/ericsson-mobility-report-november-2021.pdf

[39] A. Visagie, D. W. Swanepoel, and R. H. Eikelboom, "Accuracy of remote hearing assessment in a rural community," *Telemedicine e-Health*, vol. 21, no. 11, pp. 930–937, Nov. 2015.

[40] L. Xiao, B. Zou, L. Gao, M. Weng, M. Lando, A. E. Smith, W. Barber, and H. Yao, "A novel tablet-based approach for hearing screening of the pediatric population, 516-patient study," *Laryngoscope*, vol. 130, no. 9, pp. 2245–2251, Sep. 2020.

[41] I. Almufarrij, H. Dillon, P. Dawes, D. R. Moore, W. Yeung, A.-P. Charalambous, C. Thodi, and K. J. Munro, "Web- and app-based tools for remote hearing assessment: A scoping review," *Int. J. Audiology*, pp. 1–14, Jun. 2022.

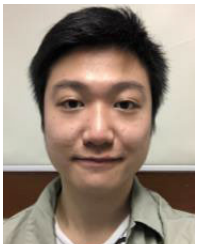
[42] I. L. Howell, K. J. Gautereaux, J. Glasner, N. Perna, C. Ballantyne, and T. Nestorova, "Preliminary report: Comparing the audio quality of classical music lessons over zoom, Microsoft teams, VoiceLessonsApp, and Apple FaceTime," Dept. Voice Pedagogy, Dept. Voice, New England Conservatory Music, Boston, MA, USA, Tech. Rep., 2020.

[43] K. Patel, L. Thibodeau, D. McCullough, E. Freeman, and I. Panahi, "Development and pilot testing of smartphone-based hearing test application," *Int. J. Environ. Res. Public Health*, vol. 18, no. 11, p. 5529, May 2021, doi: 10.3390/ijerph18115529.

[44] M. Masalski, L. Kipinski, T. Grysinski, and T. Krecicki, "Hearing tests on mobile devices: Evaluation of the reference sound level by means of biological calibration," *J. Med. Internet Res.*, vol. 18, no. 5, p. e130, May 2016, doi: 10.2196/jmir.4987.

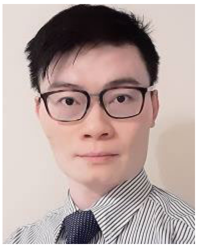
[45] G. Kim and W. Han, "Sound pressure levels generated at risk volume steps of portable listening devices: Types of smartphone and genres of music," *BMC Public Health*, vol. 18, no. 1, Dec. 2018, doi: 10.1186/s12889-018-5399-4.

[46] R. Steinmetz, "Human perception of jitter and media synchronization," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 1, pp. 61–72, Jan. 1996.



TAK CHUN KWONG (Student Member, IEEE) received the B.Eng. degree in mechanical engineering from The Hong Kong Polytechnic University, where he is currently pursuing the Ph.D. degree with the Department of Mechanical Engineering. His research interests include signal processing and environmental noise control.

YAT SZE CHOY received the B.Eng. degree (Hons.) from the Department of Building Services Engineering, The Hong Kong Polytechnic University, and the Ph.D. degree from the Department of Mechanical Engineering, The Hong Kong Polytechnic University, in 2003. She was a Research Assistant and a Research Associate with the Department of Mechanical Engineering, in 2003 and 2004, respectively, where she researched aeroacoustics. She was a Lecturer with the Hong Kong Institution of Vocational Education, Tsing Yi, in 2006. She joined The Hong Kong Polytechnic University, in 2007, where she is currently an Associate Professor. Her research interests include sound and vibration, building acoustics, environmental and duct noise control, sound source identification technique, and physiological and psychological acoustics. She is the Guest Editor of *Acoustics* (MDPI) and an Editorial Board Member of *Applied Science*.



STEVE W. Y. MUNG (Senior Member, IEEE) received the B.Sc. degree (Hons.) in applied physics from the City University of Hong Kong (CityU), in 2006, and the M.Phil. and Ph.D. degrees from the Department of Electronic Engineering, CityU, in 2009 and 2012, respectively.

He has more than 20 years of industrial experience in operations, design, and manufacturing from semiconductors to consumer products. He is currently a Research Assistant Professor with the Research and Development Office, The Education University of Hong Kong (EdUHK), and an Adjunct Associate Professor with the Department of Electrical and Electronic Engineering, The University of Hong Kong (HKU), and the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University (PolyU). He is the Co-Founder and the Company Representative of the PolyU-Innovation Technology (HK) Audio and Speech Signal Processing Joint Laboratory. He has published and invented more than 100 journal articles, conference papers, and China/U.S. patents related to wireless/speech/audio technologies. His research interests include microwave/millimeter-wave integrated circuits and devices, antenna design, the IoT applications, and wireless communications.

Dr. Mung is a fellow of the Institution of Engineering and Technology, the Institute of Physics, and the Institute of Science and Technology. He is a member of the Reviewer's Board in the MTT-Sat Challenge. He serves as a Technical Committee Member for the MTT-S TC-4 Microwave Passive Components and Transmission Line Structures Committee, the Audio/Video Systems and Signal Processing/Smartphone and Mobile Device Technologies Technical Committee in the IEEE Consumer Technology Society, and the Instrument Science and Technology Group Committee in the Institute of Physics. He is also an Associate Editor of *IEEE Consumer Electronics Magazine*. He serves as a reviewer for several IEEE journals. He is a Chartered Engineer with the Engineering Council, U.K. He is a Chartered Scientist with the Science Council, U.K.

• • •