TAPIR Sound Tag: An Enhanced Sonic Communication Framework for Audience Participatory Performance

Jimin Jeon
Graduate School of Culture Technology, KAIST
Daejeon, Korea
jimin.jeon@kaist.ac.kr

Gunho Chae
Graduate School of Culture Technology, KAIST
Daejeon, Korea
chaegunho@kaist.ac.kr

Edward Jangwon Lee
Graduate School of Culture Technology, KAIST
Daejeon, Korea
noshell@kaist.ac.kr

Woon Seung Yeo
Division of Digital Media
Ewha Womans University
Seoul, Korea
woony@ewha.ac.kr

ABSTRACT

This paper presents an enhanced sonic data communication method using TAPIR (Theoretically Audible, but Practically Inaudible Range: frequencies above 18kHz) sound and a software toolkit as its implementation. Using inaudible sound as a data medium, a digital data network among the audience and performer can be easily built with microphones and speakers, without requiring any additional hardware.

"TAPIR Sound Tag" is a smart device framework for inaudible data communication that can be easily embedded in audience participatory performances and interactive arts, by enabling peer-to-peer or broadcasting real-time data communication among smart devices. The developed toolkit can be used without any advanced knowledge in signal processing and communication system theory; simply specifying carrier frequency and bandwidth with a few lines of code can start data communication. Several usage scenarios of the system are also presented, such as participating in an interactive performance by adding and controlling sound, and collaborative completion of an artist's work by audience.

We expect this framework to provide a new way of audience interaction to artists, as well as further promoting audience participation by simplifying the process: using personal smart devices as a medium and not requiring additional hardware or complex settings.

Keywords

Audience Participation, TAPIR sound, Sonic Data Communication

1. INTRODUCTION

Over the past few decades, there have been myriad attempts to apply personalized interfaces on collaborative music making and interactive performance that the audience can participate in. After the introduction of smart devices, such as smart-phones and tablets, encourage both the performer and audience to use their own devices as an interface. This has brought a paradigm shift on participatory music making and performance[11, 5].

With new smart devices with sensing and networking features such as iPhone, various types of data and signals can be easily collected from audience activities or the physical environment without separate sensor devices. The smart devices act as a platform that collects data and converts them to the materials for interactive musical expression instantly[3]. Using smart devices as the personalized interfaces lightens the burden of designing additional hardware on artists, thus allows them to fully concentrate on content production.

However, since technologies such as WiFi is a common prerequisite for communication among smart devices, audiences who want to participate in the performance are required additional configurations such as joining networks. This has acted as a barrier, reducing user accessibility[8].

This research analyzes the constraints of using wireless communication in participatory performance and collaborative music creation, while proposing a 'Sound tagging system' based on "Theoretically Audible, but Practically Inaudible Range (TAPIR)" sound[15] — typically higher than 18 kHz — as an alternative communication method for participatory performance and its software framework. The framework is expected to help users without advanced knowledge in signal processing to utilize 'TAPIR Sound' for their new musical expressions. Additionally, This paper introduces scenarios of participatory performance using the TAPIR Sound Tagging system.

2. BACKGROUND

2.1 Data Communication System for Mobile Devices in Participatory Performances

Personalized interfaces for audience participation in interactive performances usually require wireless data transmission technologies among them. When using smart devices, Open Sound Control (OSC) alongside with local area network (LAN) is commonly used. This setting is known to work stably in most environments and allows users to receive and transmit data real-time[13, 2].

However, the audience must join a specific network and undergo additional settings in order to participate in the performance. This sequence acts as a great barrier in performance participation, which decreases user accessibility. Various techniques have been developed and used to solve this problem, such as Bonjour which is a zero-configuration

NIME'14, June 30 – July 03, 2014, Goldsmiths, University of London, UK.

networking protocol,[9] and state of the art web technologies using JavaScript and HTML5[12]. However, the inherent problem when using the computer network was still not solved, such as the requirement of additional equipment (Server and Router) for network system construction. This results in an increased cost corresponding to the number of participants.

2.2 Using Sound as a Medium

Many attempts to use sound instead of radio waves as a new medium of data transmission system for various purposes can be found throughout the literature. "Chirp"[1], a mobile application, uses sound from mobile devices to transmit a URL to the user's uploaded content.

There are many benefits from transmitting data signal through air, as a form of sound. By encoding data signals into audio, it is possible to transmit information to users' device by using built-in microphones, without connecting the same LAN. Since connection to the Internet or a local router is not required, no additional equipment is needed for the formation of the service. Furthermore, signal broadcasting is also possible, in addition to peer-to-peer communication. Hence, an ad-hoc network, with a single transmitter and multiple receivers, can be formed immediately. Therefore, using existing infrastructure such as stationary loudspeakers in halls, this broadcast network can be easily constructed and expanded without additional costs.

In addition, by deliberately hiding the communication process, these systems are able to provide an additional route for transmitting assistant information without disturbing auditory experience. Especially, "Theoretically Audible, but Practically Inaudible Range" (TAPIR) sound has value in this aspect. Theoretically, the hearing range (audible frequency band) of human auditory system (HAS) is from 20Hz to 20kHz. However, in reality, in reality, the majority of people cannot hear sound on 18kHz or a higher band. In contrast, most audio devices are designed to have a capability of accepting audio up to 22kHz – 24kHz. That is, a marginal frequency band exists which humans cannot recognize but machines can. This audio band can be utilized for various purposes, especially for data communication without interfering the listener's cognition.

"Shopkick"[10] is a successful example of using inaudible sound for commercial purposes. In the shop, the speaker in a specific region of the store emits an audio signal, which is mapped to the store's unique ID. When a user's smart device receives this ID signal, it automatically recognizes the corresponding shop and provides sales information. Another case is the "Yamaha Infosound"[14], which provides additional information to the audience by embedding codes in the audio contents of TV broadcasting. There are also artistic applications such as "Cryptone"[4], which implemented interactive personal presentation in crowded places.

2.3 Limitations of Previous Sonic Data Communication Systems

Despite the advantages of using sound for data communication, challenging problems prevail. Although there is less noise in high frequency audio band compared to the band we normally hear, many factors that disturb communication exist. While the sequences of signals (sound) emitted from a speaker go through the air and be resampled by the receiver's microphone, they suffer from echo (fading), amplitude/phase distortion, and noise. These subtle, but critical corruptions of signals are unable to predict. This phenomenon tends to be intensified as the signal transmission speed increases. Thus, a method of maintaining sufficiently high speed and increasing stability is needed. Moreover, in order to limit high frequency signals to the inaudible bandwidth, other technical improvements are required, such as constraining frequency bandwidth.

In order to solve such problems, many different communication technologies have been used. Chirp uses Single-Frequency (SF) signaling, while Shopkick and Cryptone use Dual-Tone Multi-Frequency (DTMF) Signaling, which are the same techniques used for dialing in telephones. These methods are able to be robust against external effects. However, these signaling methods can be considered a triggering method, rather than data transmission. The amount and variety of data is strictly limited, compared to its high possession of bandwidth. Direct sequence spread spectrum (DSSS) is the technique that is used for Yamaha Infosound. This system sends more accurate data and requires low power consumption, however exhibits low spectral efficiency.

3. TAPIR SOUND TAG FRAMEWORK

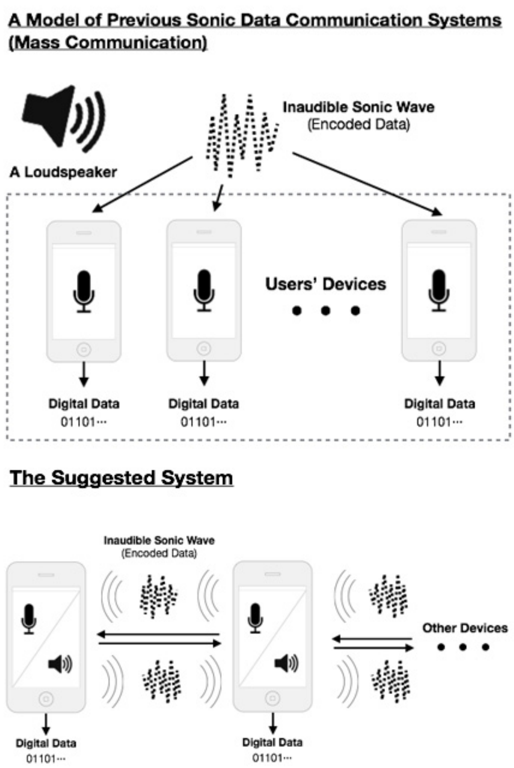

Figure 1: A System Model of Previous Sonic Data Communication (Mass Communication) and The Proposed Model

Although the research subject of data communication systems using sound is attractive, it is still in its early stage. Aforementioned systems are usually implemented with simple algorithms to retain high stability against external effects, which is exchanged with low spectral efficiency and limitations of data size. As a result, attempts to apply these systems for the purpose of intercommunication of personalized interfaces are usually limited to simple triggers, rather than data streams. Nevertheless some studies applied improved algorithms to achieve better results[7], only few have attempted to use these technologies for mutual interaction. Therefore, most of the previous systems have been limited to a form of mass communication, with one sender and multiple receivers as shown in Figure 1.

We introduce a "TAPIR Sound Tag" system in this chapter, which offers enhanced stability and a system model that allows users to simultaneously generate and analyze audio signals in real-time using their own devices. Moreover, since the system is implemented as a software platform, it can be easily applied to developing personalized interfaces in interactive artwork.

3.1 Technical Description

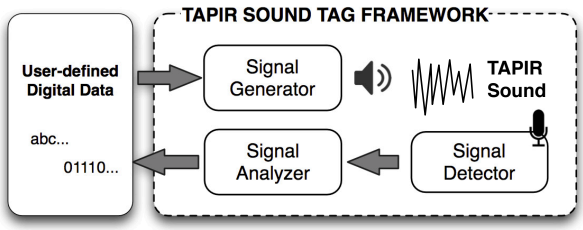

Figure 2: Schematic Diagram of TAPIR Sound Tag Framework

In the design of "TAPIR Sound Tag" system, which is a data transmission system that uses TAPIR sound having high spectral efficiency and stability, we adopted new technologies such as Orthogonal-Frequency Division Modulation (OFDM): a technique applied in cellular and digital TV networks. This puts data bytes into multiple carriers and transmits it in a parallel method, thereby maintaining a high transmittance speed while also reducing influence from surrounding environment. Furthermore, to further extend system stability, Forward Error Correction (FEC) and channel recovery techniques are applied, which are also widely used in modern radio wave communication networks. A simplified diagram of the system structure is shown in Figure 2.

In order to process high complexity mathematical operations such as filtering, FFT, and correlation etc. in real-time, especially in mobile devices with low processing power, most parts of these operations are implemented as forms of vector calculation using SIMD (Single Instruction, Multiple Data) co-processors such as SSE (for Intel x86/64) or NEON (for ARM processors). Therefore, the encoding and decoding processes of data can be done almost in real-time for the most of mobile devices. In addition, we added numerical parameters that have been obtained from experiments, which are currently considered as optimal. The result is implemented with C++ and distributed as a form of shared/static library, which can be included in applications for iOS, Android and OSX.

"TAPIR Sound Tag" system shows a high transmission rate around 200bps with a bandwidth of 900Hz, and is able to communicate at a distance of approximately five meters using ordinary portable speakers (3W output power). Under most conditions, the system is able to transmit short data streams such as URLs in a very short time (under 0.5 sec): this performance is comparable to state-of-the-art wireless tagging technologies such as NFC or RF-ID.

3.2 API Design

Our most emphasized goal is to allow artists and developers to be able to apply our framework on their projects easily, without having advanced knowledge in communication systems. For this, we have built a very simple set of

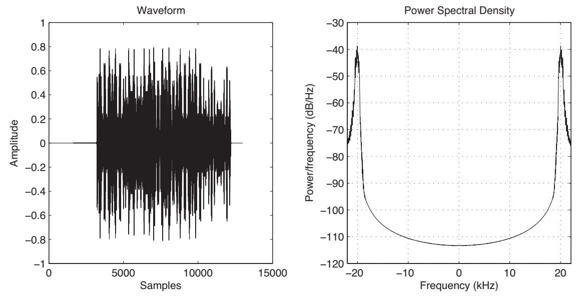

Figure 3: A Waveform and Spectral Power Density Plot of a Generated TAPIR Sound Tag Signal

APIs, where almost every procedure related to complex signal processing is automated. Thus, users of the framework can transduce a bitstream such as string to audio data in real-time with only few lines of code without suffering from complex configurations. The only required process to use our API is to specify the carrier frequency and bandwidth, just like tuning in to a radio channel. That is, similar to OSC, it is easily attachable to existing mobile applications and allows them to exchange detailed and various statuses like data from intrinsic sensors.

4. APPLICABLE SCENARIOS

As described in the previous chapter, TAPIR Sound Tag has several characteristics that enhance existing communications methods using networks such as LAN. When applied to participatory performance, collaborative music creation and interactive arts, simultaneous and prompt interaction between performer (artist) and audience or even among audiences can be formed, encouraging new artistic expression.

4.1 Participation without Barriers

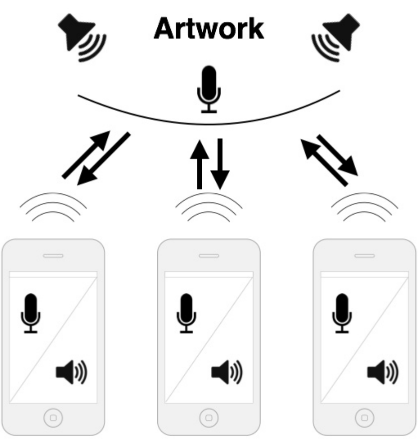

Figure 4: An Abstract Model for Audience Participatory Performances

Interactive performances and artworks that intend to induce audience participation can apply TAPIR Sound Tag to help the audience to easily participate by just playing or receiving sound through their own smart devices. This does not require additional equipment or cognitive barriers, such as configuration, which might cause the audience to hesitate in participating. Moreover, a new form of interaction can be provided, where one can immediately communicate with unidentified multiple users (Figure 4). For instance, performers can trigger devices of the audiences to show specific images or information related to the performance by playing TAPIR signal. Audiences also can reflect their intention to the interactive performance (artwork) by controlling elements of it by sending signals with their own device.

4.2 Sharing and Propagation

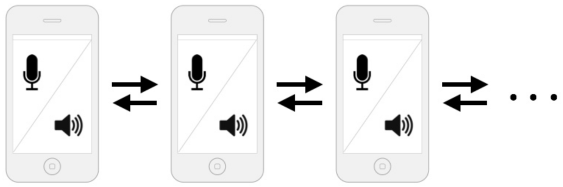

Figure 5: An Abstract Model for Sharing and Propagation

By using TAPIR Sound Tag, performers (artists) and audiences can transmit and receive data with each other by using peer-to-peer or broadcasting methods (Figure 5). Albeit the fact that TAPIR Sound Tag is a close distance communication method using speakers as a transmitter, it has an advantage of controlling the boundaries of audience participation and the presence of interactive performance (artworks) in physical space by controlling the volume of speakers. For example, TAPIR Sound tag can be a platform for generative arts/performance. Artists can share an uncompleted part of their artwork with a small number of audiences in a specific place by transmitting the code of their work through TAPIR Sound Tags. Furthermore, through the repeated data exchange between audiences, they can build their own artwork derived by the original artwork.

5. CONCLUSION AND FUTURE WORKS

In this paper, a novel sonic data communication method for creating new participatory performances and artistic expressions is proposed, using TAPIR sound. Using the framework presented through this research, performers (artists) are able to easily embed participatory interaction in their works, only using simple technology and equipment: smartphones.

Although our system shows good performance in general situations, several improvements are still being made, and there is room for much more. Our system's high spectral efficiency, which is the one of the greatest advantages, enables further utilization of the remaining TAPIR band for various purposes, such as location detection using time of arrival (ToA) of multiple signals.[6]

In conclusion, we strongly believe that the TAPIR Sound Tag framework presented in this research can also be applied to many other fields, such as broadcasting and other cultural industries. Our ultimate goal of is to apply TAPIR Sound Tag to more diverse, and appropriate fields and offer new user experience. For this purpose, we are improving our framework to be more general and flexible, to be usable in various fields.

6. REFERENCES

[1] Animal Systems. chirp.io – Let's teach the machines to sing. http://chirp.io, Feb. 2014.

[2] A. Freed and A. Schmeder. Features and future of Open Sound Control version 1.1 for NIME. In *Proceedings of the 2009 Conference on New Interfaces for Musical Expression*, Pittsburgh, USA, 2009. Carnegie Mellon University.

[3] L. Gaye, L. E. Holmquist, F. Behrendt, and A. Tanaka. Mobile Music Technology: Report on an Emerging Community. In *Proceedings of the 2006 Conference on New Interfaces for Musical Expression*, pages 22–25, Paris, France, 2006. IRCAM – Centre Pompidou.

[4] M. Hirabayashi and M. Shimizu. Cryptone: Interaction Between Performers and Audiences with Inaudible DTMF Sounds. In *SIGGRAPH Asia 2012 Emerging Technologies*, SA '12, pages 5:1–5:4, New York, NY, USA, 2012. ACM.

[5] B. Kim and W. S. Yeo. Interactive mobile music performance with digital compass. In *Proceedings of the 2012 Conference on New Interfaces for Musical Expression*, Ann Arbor, USA, 2012. University of Michigan.

[6] S. Kim, B. Kim, and W. S. Yeo. IAMHear: A Tabletop Interface with Smart Mobile Devices Using Acoustic Location. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 1521–1526, New York, NY, USA, 2013. ACM.

[7] H. Matsuoka, Y. Nakashima, and T. Yoshimura. Acoustic OFDM system and its extension. *The Visual Computer*, 25(1):3–12, 2009.

[8] J. Oh and G. Wang. Audience-participation Techniques Based on Social Mobile Computing. In *Proceedings of the International Computer Music Conference 2011*, pages 665–672, Huddersfield, UK, 2011. University of Huddersfield.

[9] C. Roberts, G. Wakefield, and M. Wright. Mobile Controls On-The-Fly: An Abstraction for Distributed NIMEs. In *Proceedings of the 2012 Conference on New Interfaces for Musical Expression*, volume 2012, Ann Arbor, USA, 2012. University of Michigan.

[10] shopkick, Inc. shopkick. http://www.shopkick.com, Feb. 2014.

[11] A. Tanaka. Mobile Music Making. In *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, pages 154–156, Singapore, 2004. National University of Singapore.

[12] N. Weitzner, J. Freeman, S. Garrett, and Y. Chen. massMobile–an audience participation framework. 2012, 2012.

[13] M. Wright and A. Freed. Open Sound Control: A New Protocol for Communicating with Sound Synthesizers. In *International Computer Music Conference*, pages 101–104, Thessaloniki, Hellas, 1997. International Computer Music Association.

[14] Yamaha Corporation. INFOSOUND. http://jp.yamaha.com/products/soundsignage/infosound/, Feb. 2014.

[15] W. S. Yeo, K. Kim, S. Kim, and J. Lee. TAPIR Sound as a New Medium for Music. *Leonardo Music Journal 22*, 2012.