immaterial.cloud: Using peer-to-peer technologies for music

Tate Carson
Louisiana State University
tcarso2@lsu.edu

ABSTRACT

immaterial.cloud is an immersive audiovisual installation that explores a possible networked future of peer-to-peer technologies, away from cloud computing. Participants experience the work via two to four smartphones placed in different locations in a room. As participants walk up to a phone, they see a representation of themselves through data. If the participant gets close enough, the phone triggers a change in the sound of *immaterial.cloud* and the other phones follow.

1. INTRODUCTION

immaterial.cloud is a web application that uses peer-to-peer technologies to send data between phones without the need for an intermediary server. *immaterial.cloud* creates the chance for a shared space with participants by using technology collectively. It requires the phones to act together, not as individuals as is usual in this era of personalized devices. Experiencing *immaterial.cloud* presents an opportunity for a restoration of attention fatigued by an overuse of technology.

The idea for *immaterial.cloud* emerged to create a communal experience during the COVID-19 shelter-in-place orders during which many people have turned to the internet for communication and entertainment. While social networks such as Facebook or Google seek our attention for profit, *immaterial.cloud* seeks a deep attention that creates a shared sense of place and time for the participants.

2. MOTIVATION

immaterial.cloud uses a peer-to-peer instead of centralized topography for networking communication. Peer-to-peer systems in computer music are already in widespread use through ad hoc networks and Open Sound Control (OSC). But both solutions have issues that prevent seamless audience participation such as a lack of a Domain Name Server (DNS) and data persistence. These systems work over local networks and work for smaller groups of people but aren't sufficient for internet communication.

The goal of peer-to-peer networking is to make setup of the system easy and seamless but without having to run a server backend. While the interaction possibilities created by networked smartphones has been worthwhile in creating a sense of interconnection, researchers should examine the infrastructure used behind these systems and explore alternatives. The aim is to develop technology that doesn't fatigue our attention or treat us as an economic good.

2.1 Attention

Why this focus on avoiding the major platforms? Because of the attention economy, which turns our attention into an economic good by measuring it through clicks, downloads, and likes. Measurement allows companies to sell our attention as data to advertisers [29], which becomes an issue because our attention is a scarce resource. As Crogan and Kinsley state, the point of paying attention to attention is to, "reanimate the potential for a less poisonous adoption of the widely recognized potential of digital audiovisual culture in order to re-form (that is, re-mediate) culture, sociality, economy and ecology today [6]."

How can the attention economy be resisted? We must redesign the technologies that encourage a capitalist perception of time, place, self, and community. Deepening one's concept of place will extend an awareness of history and current connection to everything around them. Odell suggests that "doing nothing" moves our attention from economic concerns to the physical, place based domain [18]. Doing nothing, or sitting out the attention economy is an active proposition and "entails an active process of listening that seeks out the effects of racial, environmental, and economic injustice and brings about real change [18]."

Because of these platforms' flaws we shouldn't be using them to distribute and make our art. These platforms reward a way of being that isn't beneficial to society. They resist the concept of interconnection, by emphasizing our differences and not allowing us to show fuller versions of ourselves. We must "stand apart"[18] as Odell suggests, and create new systems for communication on the internet.

How can we regain attention? One way is through art, another through nature. Conceptual art encourages a perception of time on different scales and tempos and allows for deep attention. Interacting with nature can also restore our attention. Psychologist Kaplan found that the natural environment aids in stress reduction by providing a "restorative environment" that reduces the fatigue caused by directed attention [9]. Attention Restoration Theory (ART) claims that we use directed attention in our daily lives to function. Sustained directed attention leads to cognitive depletion and mental fatigue [10]. Kaplan didn't mention sound directly

Web Audio Conference WAC-2021, July 5–7, 2021, Barcelona, Spain.

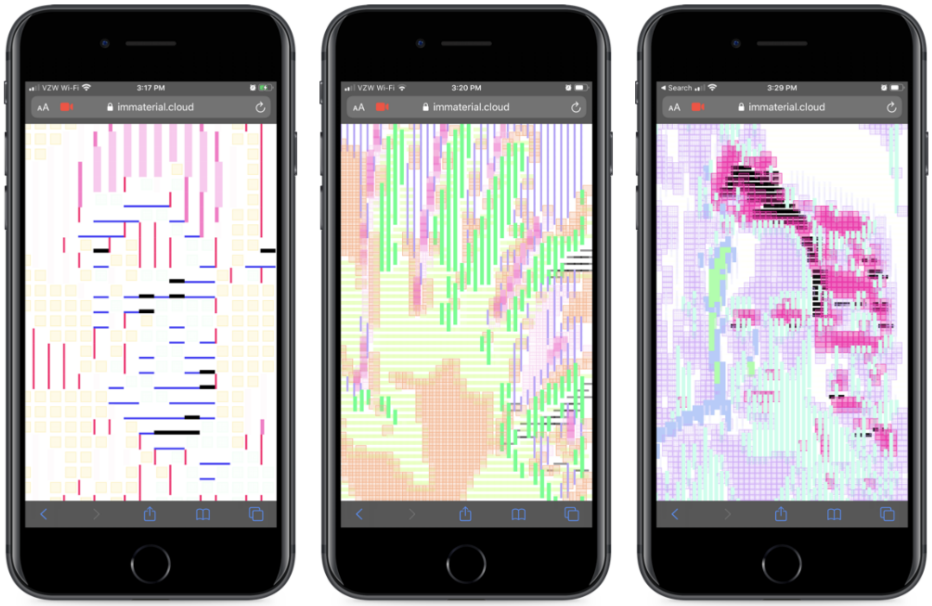

Figure 1: Visualization of user interaction on the phone

in his research, but a recent study by Ratcliffe has extended Kaplan's research to show that certain bird sounds may provide restorative benefits [21].

3. BACKGROUND

immaterial.cloud uses the audience as a speaker array, a practice with a long history in music [28]. Many frameworks and compositions use distributed smartphone speaker systems with the Web Audio API. Frameworks such as Schnell and Robaszkiewicz's Soundworks [22], Jesse Allison's Nexus [1], Weitzner and Freeman's massMobile [30] and Piquemal's Rhizome [25] are designed for artists who don't want to create bespoke systems. These frameworks run on a variety of centralized server configurations, anywhere that Node.js can run. These systems have taken a variety of networking approaches that center on cloud computing, whether running a cloud-based virtual machine (VM) such as Amazon's EC2 or Google Cloud [2], a "serverless" host like Zeit's Now [5], a Platform as a Service (PaaS) [7] such as PubNub, or server running on one's own laptop [14]. Many papers don't describe projects' hosting configurations, but because of the frameworks used, one can assume that they run on some version of the aforementioned technology [25, 8, 13, 26]. The lack of information about hosting configurations often leaves room for inquiry. What these systems have in common is that they use centralized networking topology.

Peer-to-peer systems aren't new but had their heyday in the public imagination during the height of online file sharing when Napster was operating between 1999 and 2001 [19]. Because of the popularity and common knowledge of those systems, interactive music works began to emerge that used and extended similar technologies. In 2004 Tanaka created a collaborative music-making system based on personal digital assistants (PDAs), which extend the simple sharing of music on peer-to-peer systems allowing users who were near each other to interact through Service Discovery Protocol (SDP) [27]. Lee, Essl, and Mao have used SDP effectively to distribute applications to mobile phones in a mobile ad hoc network (MANET) [11]. Creators of Network Music Performance (NMP) systems have a parallel interest in peer-to-peer architectures because they allow real-time performance in different geographic locations [24]. Because one goal of these systems is to enable real-time music performance with acoustic instruments, engineers emphasize time synchronization over ease of use or audience participation.

Peer-to-peer systems continued to develop with the creation of web standards such as WebRTC, which is currently in wide use for video conferencing. WebRTC has an underutilized capability, the data channel, which is most useful for web audio projects that might want to keep data synchronized. Even so, the data channel's use in musical projects is limited, with many projects preferring to run a server and use WebSockets.

Several web audio projects have used WebRTC. Lind's Soundtrap uses WebRTC's video channel [12]. Ramsay and Paradiso's GroupLoop [20] uses the audio channel to send audio between clients to create a feedback based performance system, and Black's Hear-Here [3] uses it to coordinate an FM radio broadcast out of the audio of distributed clients. Xing, Ulrich and Diab's Fun With Chords [31], a distributed music player, uses the data channel. Authors gave few reasons for choosing WebRTC over other methods in the cited papers, but one can assume they chose WebRTC for the reasons same that interest me—the minimal need for servers.

Researchers have proposed several peer-to-peer systems that use Open Sound Control (OSC). Roberts, Wakefield and Wright present an extension to Robert's Control program that simplifies including mobile devices in a computer music workflow [23]. Another example is GoOSC [4] by Cabrera, which emphasizes the peer-to-peer nature of the network.

4. TECHNICAL DESIGN

4.1 WEBRTC

immaterial.cloud[1] is networked using WebRTC (Web Real-Time Communication), which allows for sites to stream audio, video, or data between each other without an intermediary. WebRTC consists of several APIs and protocols that allow for peer-to-peer communication with no need for browser plugins or third-party software [17]. The protocols that are important for enabling peer-to-peer connections are ICE (Internet Connectivity Establishment), STUN (Session Transversal Utilities for NAT), and TURN (Transversal Using Relays around NAT) [15]. While WebRTC could work without servers in a perfect environment, often it needs STUN and TURN to navigate firewalls. The STUN server, also known as signaling server, performs a "handshake" between two peers that decides if the peer is reachable. After the handshake, the connection is truly peer-to-peer. If this simpler route isn't achieved, a TURN server relays packets to and from peers to traverse a NAT. These intermediaries aren't ideal if the goal is creating a mode of communication without servers, but for now they're necessary to enable the wider adoption of WebRTC and other peer-to-peer technologies.

immaterial.cloud uses the data channel capability of WebRTC, defined by the RTCDataChannel interface. This interface allows for bidirectional transfer of peer-to-peer data [16]. The work uses the PeerJS[2] library to simplify the process of setting up connections and dealing with messaging. PeerJS provides an API like Socket.io[3] to send and receive messages with few lines of code. PeerJS also includes a signaling server which allows a programmer to use the library without having to set up and run a server. Though *immaterial.cloud* uses video as it's main mode of user interaction, the RTCPeerConnection API isn't used because video isn't streamed between devices.

4.2 User Interaction

immaterial.cloud needs two to four smartphones (iPhone or Android) connected to the internet via WiFi or a cellular network. All sound during the installation is played through the phones via a web browser. *immaterial.cloud* will work with a group of participants or just one. A user interacts with *immaterial.cloud* through motion, waving a hand or walking up to a phone (see Figure 1). The camera of the phone tracks any change in motion, triggers an update to the sound playing, and updates all other phones.

A participant joins the network by going to https://immaterial.cloud and entering the ID of a chosen "host" phone. Though the "host" phone doesn't need to enter any ID to play, the participant still needs to press join to start the sound (see Figure 2). It's necessary to use IDs so *immaterial.cloud* knows which phones to send messages to (without participants having to sign up for accounts).

4.2.1 Data Transfer

immaterial.cloud uses a broadcast system to connect all peers through the server phone because WebRTC only allows communication between two peers. This server phone isn't a true server but relays messages between the nodes of the network through the hub of the server phone and back out

Figure 2: The welcome screen and the ID screen

to the other nodes. Each phone keeps a list of client IDs belonging to the other peers they're connected to.

First, the server phone comes online. Nothing is special about this phone; any participant's phone can be the server phone. The other phones connect to the server phone by entering the server phone's client ID. This step allows for each phone to keep an updated list of the participating phones' IDs in the network. Each time a phone connects to the server phone, a broadcast message is sent to every other phone, notifying them that a new phone has been added.

4.2.2 Composition Design

When each phone starts *immaterial.cloud*, it is assigned one of four presets that define its sound. I created these presets using the dat.gui[4] library, which provides a GUI for changing variables in JavaScript. Its most useful function is to save unique combinations of settings as presets to be recalled later. *immaterial.cloud* uses granular-js[5], a granular synthesis library for JavaScript created by Philipp Fromme. Granular-js makes granular synthesis with the Web Audio API much easier providing access to the density, spread, and pitch parameters of each granular sample instance.

A typical interaction might go as follows. A user waves their hand over a phone, which is assigned a preset of "deeper," then the preset is sent to each of the other phones, first traveling through the server phone to change its preset to "deeper" (see Figure 3). Figure 4 shows how data flows from the triggered phone to the server phone, then out to the other peer phones, setting the other phones sound to the chosen preset. The user doesn't have fine grained control over how the sounds develope but can explore the different presets by triggering each phone.

[1]https://github.com/tatecarson/immaterial.cloud
[2]https://peerjs.com/
[3]https://socket.io/
[4]https://github.com/dataarts/dat.gui
[5]https://github.com/philippfromme/granular-js

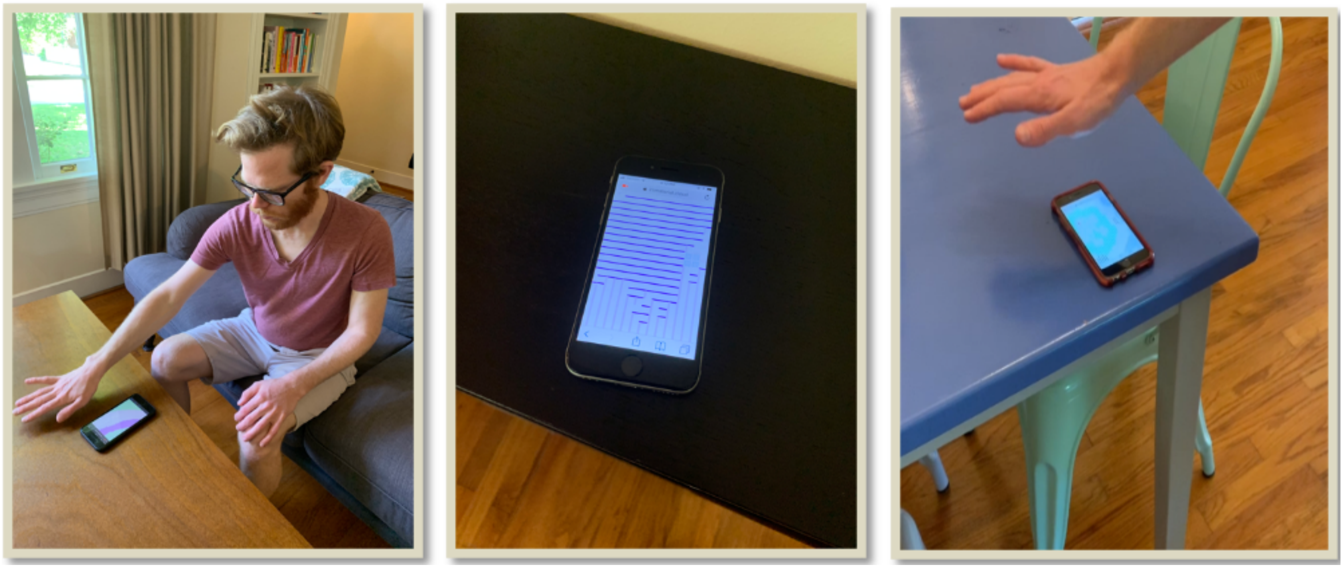

Figure 3: Participant waving to interact with the phone

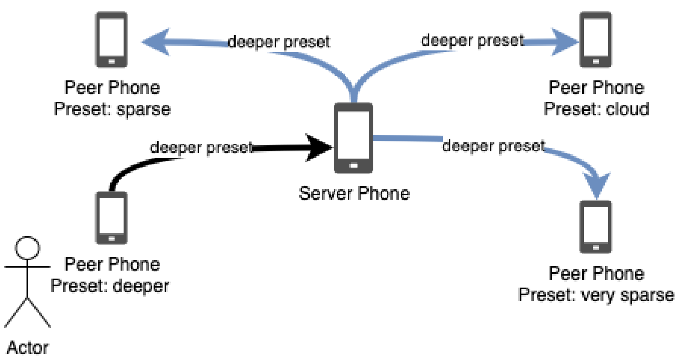

Figure 4: Sending a preset from one peer to the others

Each preset uses one of four granulated samples, a ring tone, a double bass improvisation, a music box, or tubular bells. The presets control the pitch, attack, release, and density for each grain of the granular synthesis. When a player triggers a new preset, the system interpolates between the previous preset and the new preset over a randomly chosen length of time. This combination of presets and interpolation provides different sound possibilities without too much chaotic variation.

5. CONCLUSIONS AND FUTURE WORK

Peer-to-peer technologies provide a pathway to a form of networked communication that doesn't rely on the attention economy. *immaterial.cloud* is an early step in the movement towards peer-to-peer networked art. I focused on the attention economy to describe why it is important to find alterative means for networking. It's important for artists to explore technologies independent from major platforms to ensure control over their work.

While *immaterial.cloud* uses WebRTC, I think it is important to outline possible future protocols. Many technologies are currently in development that seek to invert the current client/server architecture of the web. These systems share a vision for a decentralized internet. Some popular protocols include Interplanitary File System (IPFS), Secure Scuttlebutt[6], Hypercore Protocol[7], and Matrix[8]. These protocols vary in design and goals. One could imagine a composer choosing a specific protocol as a compositional choice that would make for more interesting modes of interaction for our networked web audio creations of the future.

6. REFERENCES

[1] J. Allison, Y. Oh, and B. Taylor. NEXUS: Collaborative Performance for the Masses, Handling Instrument Interface Distribution Through the Web. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pages 1–6, 2013.

[2] J. Allison, F. Ostrenko, and V. Cellucci. Active server roles for extended distributed performance complexity in diamonds in dystopia. In *Web Audio Conference*, London, UK, 2017.

[3] A. Black. Hear-Here: A choreographed peer-to-peer network for live participation on the radio. In *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*, pages 1–6, London, UK, 2017.

[4] A. Cabrera. Serverless and peer-to-peer distributed interfaces for musical control. In *NIME*, pages 355–358, 2015.

[5] T. Carson. Mesh Garden: A creative-based musical game for participatory musical performance. In *New Interfaces for Musical Expression*, pages 339–342, Porto Alegre, Brazil, 2019.

[6] P. Crogan and S. Kinsley. Paying attention: Towards a critique of the attention economy. *Culture Machine*, 13, 2012.

[6] https://scuttlebutt.nz/
[7] https://hypercore-protocol.org/
[8] https://matrix.org/

[7] A. Deusany, S. W. Lee, H. Ave, A. Arbor, G. Essl, H. Ave, and A. Arbor. Understanding Cloud Service in the Audience Participation Music Performance of Crowd in C[loud]. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 176–181, Brisbane, Australia, 2016.

[8] B. Houge and J. Sánchez. The Tomb of the Grammarian Lysias: Real-Time Performance and Crowd-Distributed Music Diffusion with Networked Mobile Devices. In *The 21st International Symposium on Electronic Art, Simon Fraser University, Vancouver, BC, Canada*, 2015.

[9] S. Kaplan. The restorative benefits of nature: Toward an integrative framework. *In: Journal of Environmental Psychology.*, 15:169–182, 1995.

[10] S. Kaplan and R. Kaplan. The visual environment: Public participation in design and planning. *Journal of Social Issues*, 45(1):59–86, 1989.

[11] S. W. Lee, G. Essl, and Z. M. Mao. Distributing Mobile Music Applications for Audience Participation Using Mobile Ad-hoc Network (MANET). In *NIME*, pages 533–536, 2014.

[12] F. Lind and A. MacPherson. Soundtrap: A collaborative music studio with Web Audio. London, UK, 2017.

[13] N. Madhavan and J. Snyder. Constellation: A musical exploration of phone-based audience interaction roles. Atlanta, GA, USA, 2016.

[14] A. T. Marasco and J. Allison. SoundSling: A Framework for Using Creative Motion Data to Pan Audio Across a Mobile Device Speaker Array. In *Web Audio Conference*, Berlin, Germany, 2018.

[15] MDN Contributors. Introduction to WebRTC protocols. https://developer.mozilla.org/en-US/docs/Web/API/WebRTC_API/Protocols, July 2019.

[16] MDN Contributors. RTCDataChannel. https://developer.mozilla.org/en-US/docs/Web/API/RTCDataChannel, Apr. 2020.

[17] MDN Contributors. WebRTC API - Web APIs | MDN. https://developer.mozilla.org/en-US/docs/Web/API/WebRTC_API, June 2020.

[18] J. Odell. *How to Do Nothing: Resisting the Attention Economy.* Melville House, Brooklyn, London, 2019.

[19] A. Oram. *Peer-to-Peer: Harnessing the Power of Disruptive Technologies.* " O'Reilly Media, Inc.", 2001.

[20] D. B. Ramsay and J. A. Paradiso. Grouploop: A collaborative, network-enabled audio feedback instrument. In *New Interfaces for Musical Expression*, pages 251–254, Baton Rouge, Louisiana, 2015.

[21] E. Ratcliffe, B. Gatersleben, and P. T. Sowden. Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology*, 36:221–228, 2013.

[22] S. Robaszkiewicz and N. Schnell. Soundworks–A playground for artists and developers to create collaborative mobile web performances. In *WAC-1st Web Audio Conference*, 2015.

[23] C. Roberts, G. Wakefield, and M. Wright. Mobile Controls On-The-Fly: An Abstraction for Distributed NIMEs. In *NIME*, 2012.

[24] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti. An overview on networked music performance technologies. *IEEE Access*, 4:8823–8843, 2016.

[25] T. Shaw, S. Piquemal, and J. Bowers. Fields: An Exploration into the use of Mobile Devices as a Medium for Sound Diffusion. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 281–284, Baton Rouge, Louisiana, 2015.

[26] A. Stolfi, M. Barthet, F. Goródscy, A. Deusany, and F. Iazzetta. Open band: Audience creative participation using web audio synthesis. 2017.

[27] A. Tanaka. Mobile music making. In *Proceedings of the 2004 conference on New interfaces for musical expression*, pages 154–156, 2004.

[28] B. Taylor. A History of the Audience as a Speaker Array. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, volume 481-486, Aalborg University Copenhagen, Denmark, 2017.

[29] T. Terranova. Attention, economy and the brain. *Culture Machine*, 13, 2012.

[30] N. Weitzner, J. Freeman, Y.-L. Chen, and S. Garrett. massMobile: Towards a flexible framework for large-scale participatory collaborations in live performances. *Organised Sound*, 18(1):30–42, 2013.

[31] L. Xing, M. Ulrich, and O. Diab. Fun With Chords–A Distributed Music Player.