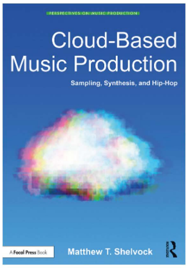
CHAPTER SAMPLER

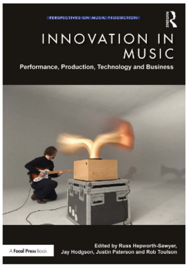
Innovations in Audio

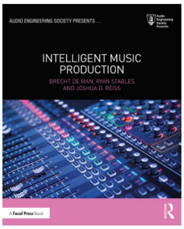
Routledge
Taylor & Francis Group

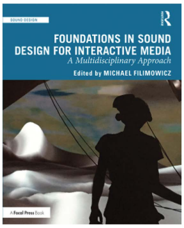
www.routledge.com

Contents


Cloud-Based
Music Production
Sampling, Synthesis, and Hip-Hop
Matthew T. Shelvock


INNOVATION IN
MUSIC
Performance, Production, Technology and Business
Edited by Russ Hepworth-Sawyer,
Jay Hodgson, Justin Paterson and Rob Toulson


INTELLIGENT MUSIC
PRODUCTION


FOUNDATIONS IN SOUND
DESIGN FOR INTERACTIVE MEDIA
A Multidisciplinary Approach
Edited by MICHAEL FILIMOWICZ

Introduction

Cloud-Based Music Production (CBMP)

This book covers the topic of cloud-based music production (CBMP) tools as they are used within today's hip-hop beat-making scene. A host of these CBMP services, which support the creation of sample-based and synthesizer-based music, have recently become widespread in the music production community. As a result, this book focuses on key CBMP features which shape the way that music producers work. Some of these features include easily searchable sample libraries, which contain sounds made by professional producers; a host of Digital Audio Workstation (DAW) plugins, such as synthesizer VSTs and mixing plugins (as well as presets for these tools); and, in some cases, these services can offer free recording software (i.e., DAWs), collaboration tools, contests, and other forums for online community engagement. Throughout this book, I explain CBMP music-making methods using both text and audio examples. Readers are encouraged to listen to these audio files, which can be found on this book's webpage on the Routledge website (www.routledge.com/9780815353195).

CBMP services are extremely popular in today's music-making communities, and, at the time of writing, one platform known as Splice boasts over 2.5 million subscribers (King 2019). This means that over 2.5 million music makers are currently *creating* music using only one of the available CBMP platforms. Several other well-known services exist, such as Loopcloud, Noiiz, and 99 sounds. Thus, the CBMP community accounts for a large portion of the current music-making population, and, in fact, CBMP skills have become a new standard competency for many contemporary producers. Although creatives working in numerous genres use CBMP tools, these technologies have been particularly appealing for hip-hop producers who, historically, were known to sift through crates of vinyl records in order to find audio samples from which they could create beats. Contemporary beat makers have since added a number of new techniques to their repertoire because of advances in digital recording tools and cloud-computing. This book clarifies the ways that CBMP tools support the creation of hip-hop beats in more current studio settings, such as modern project studios.

The topic of hip-hop production also has a special relevance at the moment since hip-hop is extremely popular for music fans. One out of every three music streams in the USA in 2018 were from hip-hop records (Resnikoff 2019). In the same year, streaming services amassed a total of 901 billion total streams, meaning that hip-hop streams accounted for approximately 303.3 billion streams (Caulfield 2019). As a result of this recent boom in hip-hop streaming, this book focuses on current (c. 2009–2019) personal computer-based hip-hop music and production techniques which largely support the creation of this new music, rather than on the DJ-centered production styles of the classic hip-hop era (c. 1980–1999). However, in some cases throughout this book where these older techniques are relevant, I explain the influence of classic hip-hop sounds on today's hip-hop production scene since, as readers will learn, vintage hip-hop sounds remain sonically influential as the genre continues to transform.

0.1 DEFINING MUSIC PRODUCTION

For both scholars and musicians, confusion abounds on the topic of music production. And who could blame them? In the past, professional record creation occurred almost exclusively in expensive private studios, and, apart from those directly involved on a given recording project, few were able to witness the actions of legendary music producers such as Phil Spector in the 1960s, for example. In other words, historically speaking, music production was a more private affair, and, as a result, very few people had access to studios and recording equipment.

However, in the 1980s, this changed when recording studios began to incorporate digital technologies, and, by the late 1990s, music production equipment became much more readily available to all musicians. Cumbersome analog processes which required bulky hardware were digitally mimicked—large (and expensive) racks of analog gear were now accessible through a computer screen. Large groups of musicians took note, and, as a result, they began to create their own records. In gaining access to recording tools, home (and project)-based recordists could now easily execute standard record production techniques; for many of these producers, the home studio became a gateway to the professional music industry. As a result of this increased access to equipment, amateurs and semi-professionals were now able to participate in music production *en masse*, and, at the same time, recording skills became a major aspect of one's musical literacy—a rite of passage to the professional music world. Very quickly, machine-based music making became a lingua franca across nearly all commercial music genres.

During this time, the music industry also began to see the of a number of independent music aggregators such as CD Baby and, later on, Soundcloud and Bandcamp. Now, musicians could easily record and release their own music. New production styles, methods, and sounds began to emerge as a consequence of this digital studio boom, which began in the late 1990s and early 2000s. All of a sudden, the term *music producer* could refer to a wunderkind like George Martin—the so-called fifth Beatle—who is famous for his mastery of instrumental arrangement, songwriting, and record production skills, or it could refer to an electronic musician like

DJ Tiesto, whose command of newer musical techniques and technologies earns him a nightly gross of $250,000.[1] Today, musicians like Tiesto primarily create records from their own home studios. While home recording was once exceptional, today it is the norm.

Music production continues to change and evolve, and, as a result, the term is used in more than one way (Hepworth-Sawyer & Golding 2011: 3). However, recent formal research has caused more confusion regarding the term *production*, as many scholars who do not participate in record production define the term to mean "everything done to produce a recording of music." This overly inclusive definition does not withstand analytical scrutiny nor does it resemble (or clarify) how actual music production cultures use the term (Hepworth-Sawyer & Hodgson 2017: xii). It remains unclear how such a broad definition can benefit recordists or researchers who study actual record production activities, especially when the term already refers to a number of clearly identifiable activities.

In order to refocus how scholars use the term music production, I have suggested elsewhere that any reasonable definition of music production must consider how the term is used by music producers and others who participate in the culture of record production because, for these people, a *producer* is someone who demonstrates a specific set of musical competencies pertaining to the *production* of a record (Shelvock 2017a, 2019). These competencies differ from other types of collaborators, such as session musicians or other performers, insofar as a producer is the person who is responsible for overseeing the encoding of a musical experience within a replayable document. In fact, the operative structure of the recording industry, like any other, is wholly dependent on the contributions of skilled workers of various types. Various personnel provide the recording industry with a number of different services from business tasks concerning administration and sales to creative tasks such as production and performance. As researchers, it is important we understand these roles as they exist within actual (i.e., observable, knowable, or otherwise well-documented) music scenes.

At the outset, for some readers, production may seem like a fairly straightforward task: record producers create records. However, to make matters confusing, a record producer's skillset might include the application of a number of traditional musical competencies during the production process, such as songwriting, arranging, or composing music—so long as the producer does these activities for the purpose of creating a record. In order to create records, one has to understand how to convey a musical design—whether we call this design an arrangement, song, composition, or beat—using the medium known as recording. The creation of recordings requires familiarity with processes such as tracking, editing, mixing, and mastering; in hip-hop, sampling and synthesis are also crucial techniques.

Producers may possess any combination of the competencies discussed above, and, in turn, a producer's skillset dictates how one works. In their book *What Is Music Production*, Craig Golding and Russ Hepworth-Sawyer suggest that engineers can be broadly classified into two camps: engineer-producers and musician-producers (2011: 3–9). Musician-producers often hire trusted engineers to handle production tasks which take considerable time and energy, and they tend to focus on songwriting, arrangement, coaching musicians, and personality management. Engineer-producers, on the other hand, tend to focus on recording, mixing, synthesis, sampling, and other studio effects, and these producers may hire outside help for instrumental arranging, orchestrating, or even songwriting. However, both types of producers possess skills as both *musicians* and *engineers*, even though their practices may favour one area over the other (2011: 8–9).[2]

In hip-hop, most (if not all) producers fall into the engineer-producer category, and they make beats primarily using digital recording platforms, sample-triggering devices, pre-recorded audio (i.e., samples), and MIDI. At the same time, numerous episodes of *Rhythm Roulette* demonstrate that producers such as Rahki (Top Dawg Entertainment), Jake One (G-Unit, 50 Cent, Brother Ali, MF DOOM), and Brasstracks (Chance the Rapper) also freely incorporate instruments they play themselves, such as keys, bass, and guitar.

These sounds, whether they are collected via sampling techniques or live recording, are almost always collated within a digital audio workstation (DAW) of some type. Some popular examples for hip-hop production are Ableton, Logic, Pro Tools, and Fruity Loops, although others are used as well. During record production, the DAW becomes a partner to the producer throughout the creative process. Practically speaking, this is accomplished in two ways. When using a DAW, recordists can experiment with endless permutations of sounds, instruments, rhythms, and melodic or harmonic events, and all of these modifications can be compared, or *auditioned*, at any time. In fact, the DAW itself is considered a crucial instrument for songwriting by many, as Mark Marrington writes (2017: 77):

> Like the piano or guitar in previous eras of songwriting, [the DAW] is an instrument in its own right which impacts upon the conception and organization of musical ideas. To put it another way, the DAW has its own particular creative "paradigms" to contribute to the songwriter's process, which once understood, can be harnessed to great effect.

And, indeed, the DAW has had a remarkable impact on songwriting, arrangement, and composition, in addition to its primary functions as a tool for tracking, editing, mixing, and mastering records. This is particularly true in hip-hop, where DAWs are used as a cornerstone for beat-making.

Since hip-hop records accounted for more than 1 in 3 of all music streams in the USA in 2018, it is safe to say that the music-making techniques used to create these tracks are one of today's most relevant musical activities. Yet, as it stands, music scholars and educators know little about this process. When authors do broach issues surrounding hip-hop production, such as sampling, they typically fail to discuss the technique in a way which resembles what musicians and producers actually do in studios.

Instead of directly discussing these practices, most authors instead focus on the application of cultural or sociological theorization to the object or music production. In fact, in most cases, published accounts of sampling and production are entirely disconnected from the practice of beat-making *per se*, as it occurs within contemporary hip-hop music-making subcultures. Notwithstanding this, in rare cases wherein hip-hop beat-making receives a more methodologically rigorous treatment by authors, technological practices (i.e., production activities) are often *black boxed*, or these texts simply do not cover the act of music production *per se*. Quite often, such studies provide an excellent post hoc analysis of records, but they do not set out to directly address the techniques that created these records (Sewel 2013, 2014; Schloss 2014).

Instead, in this book, I focus my analysis on the creative inputs which *create records*, rather than the identification of these same techniques *on available records*, like so many other papers and books have already done. Readers should be aware that simply aurally identifying the use of audio samples on hip-hop records, for instance, has little to do with the analysis of how a sample was (i) captured, (ii) edited, (iii) sonically altered (i.e., mixed), and (iv) repurposed within a new composition or arrangement. As beat makers know, each of these production steps is rife with sophisticated aesthetic nuances, and a record's final sonic characteristics can dramatically change during any one of these stages (i–iv). Until researchers establish an account of these studio techniques in way which actually reflects their usage within music-making communities, we will not be able to substantively or accurately critique them.

An entire subfield is dedicated to the accurate discussion and analysis of recording practice, and more authors are beginning to discuss the practice of music production *per se*. The emergence of music production studies (MPS) in 2017, as it is called, has begun to cause scholars to evaluate the competencies required to create records. This book, of course, is aligned with the goals of this subfield. A key difference in MPS, versus previously available analyses of records and studios, is that music production scholars directly analyze *production*, or the creation of records. Quite often, these scholars possess training, or experience in executing, the music production techniques which they discuss. Since we are only now beginning to see the emergence of accurate music production research, it makes sense that researchers with experience in professional production scenes are currently leading the charge, so to speak. And, since these recordist-researchers actively make (or made) recordings, they are better positioned than other types of scholars to discuss the topic. As the emergence of MPS demonstrates, this type of experiential scholarship is on the rise within music departments.

Later in this book, I demonstrate a variety of beat production techniques and case studies. Researchers and musicians will find these cases useful for understanding how hip-hop beat producers *use* tools and techniques such as synthesis, sampling, MIDI, sample-triggers, sequencers, and signal processing devices. By providing these demonstrations of professional productions, I hope to elucidate a number of under-discussed production practices which are extremely prevalent in today's music scene. To clarify, my goal is not to cause the cultural *uplift* of these techniques in any way. Rather, the goal of this book is to prompt scholars and musicians to begin to engage with a number of musical instantiations that are as common as cadences in classical music or fretboard-tapping and power chords in heavy metal (Hodgson 2011).

0.2 CBMP SERVICES

Cloud-computing is a popular information technology systems paradigm wherein users are given access to remote networks which store and manage data, rather than a local server (or one's own computer). Cloud-based music production (CBMP) services use cloud-computing technologies to provide recordists with access to large libraries of audio samples, synthesizer presets, VSTs (including both synthesizers and mixing tools, in some cases), partially completed works, or completed works for remixing (Figure 0.1). In addition, some CBMP services such as Splice also allow users to automatically back up their work on the host's servers, as well as providing web-based beat-making tools. The most prominent cloud-based music production subscription services are Splice, Noiiz, and Loopcloud.

All of the major CBMP services offer what they refer to as *royalty-free* audio samples, which users can freely incorporate into their songs without fear of copyright litigation. In other words, these samples are free of the typical master/publishing licensing restrictions which accompany commercially available music. This term, which is commonly used in CBMP communities, describes the lack of licensing restrictions in connection with these recordings. So, users can freely incorporate these recordings

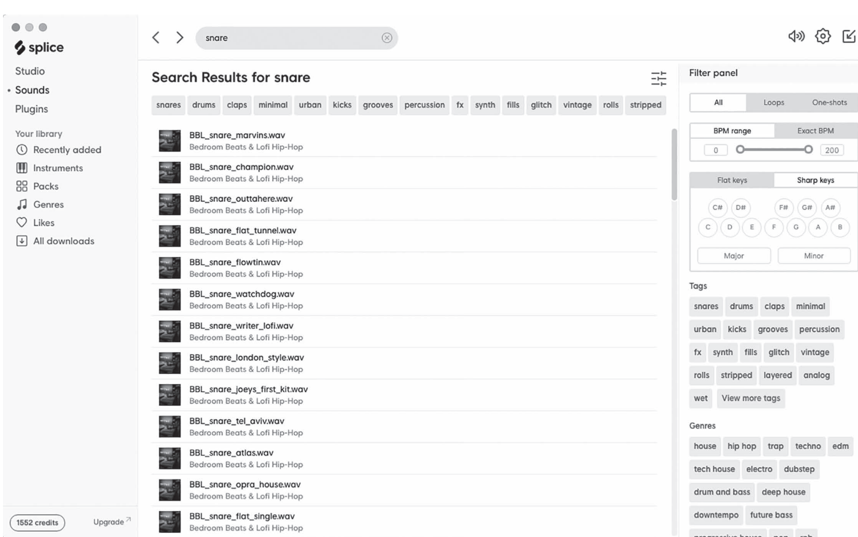


Figure 0.1 Illustrates the GUI in Splice.

(or segments from them) within their work in almost any way they would like to do so.

These *royalty-free* sample libraries receive use from a fairly wide audience. Splice, for example, has over 1 million subscribers. And, at the time of writing, 118 record labels provide Splice with numerous types of audio samples.[3] In addition, users have access to synthesizer presets. Splice co-founder Steve Martocci explains their sample library, stating (in Pangburn 2016):

> When most people think of "samples" they think of stealing a piece of audio from someone else's song and using it without their permission, but it's much wider than that. There are entire labels that fund musicians to produce samples designed to be sold in royalty free "packs". This means as long as you pay for the sample, you can use it in any of your musical work without paying any fees back to the sound designer. We've seen more and more artists making significantly more money selling samples than they make counting on their music sales or streaming revenues. These sounds range from "one shots", which could be a recording of someone hitting a snare drum, to loops which could be 30 seconds of a full drum kit. The vocal samples have also been very popular on the platform—they help give your song an organic feel. Splice Sounds also has presets files for popular virtual synthesizers like Massive and Serum.

Indeed, a popular, yet false, belief persists that sampling audio always constitutes an act of theft or social resistance of some type. While this may have been true in the past, this is not the type of sampling activity that most contemporary (c. 2009–2019) hip-hop producers primarily engage in—although it may happen on occasion. As Noah 40 (Drake, Lil Wayne, Action Bronson) explains, "I'm a fan of publishing, so I try not to sample [from copyrighted material] *too* much, but it's a tool in your arsenal."[4] In addition to Noah 40, numerous other famous producers also incorporate CBMP-based royalty-free samples within their songs. For instance, Kendrick Lamar's producer Soundwave used a famous Loopmasters sample from Splice on his 2017 album *Damn* (available until April 2017) called "COF_125_Am_LaidOut_Underwater.wav."[5]

However, royalty-free sample libraries predate the existence of CBMP services by many years. The main innovative feature of services such as Splice, Noiiz, and Loopcloud is that they aggregate numerous royalty-free sample libraries using centralized digital databases. By collecting these resources, CBMP platforms make sample-based music production much more efficient. For instance, Big Fish Audio has provided beat makers with royalty-free sample libraries since 1986.[6] Today, CBMP services provide streamlined access to samples that labels such as Big Fish Audio create.

While providing users with easy access to audio samples is their core service, many CBMP platforms also provide plugins, such as synthesizers, compressors, or EQs, as well as synthesizer presets. Splice and Noiiz, for instance, provides users with patches for Native Instruments' *Massive*, Xfer Records' *Serum*, Lennar Digital's *Sylenth*, and others. Thus, CBMP services provide users with access to networks of sounds, sound generators, and sound-shaping devices.

0.3 SOUND-TAGGING AND SEARCH FILTERS

To aid searchability, CBMP services *tag* samples and presets with descriptive language, which users can use to filter search results (Figure 0.2). For

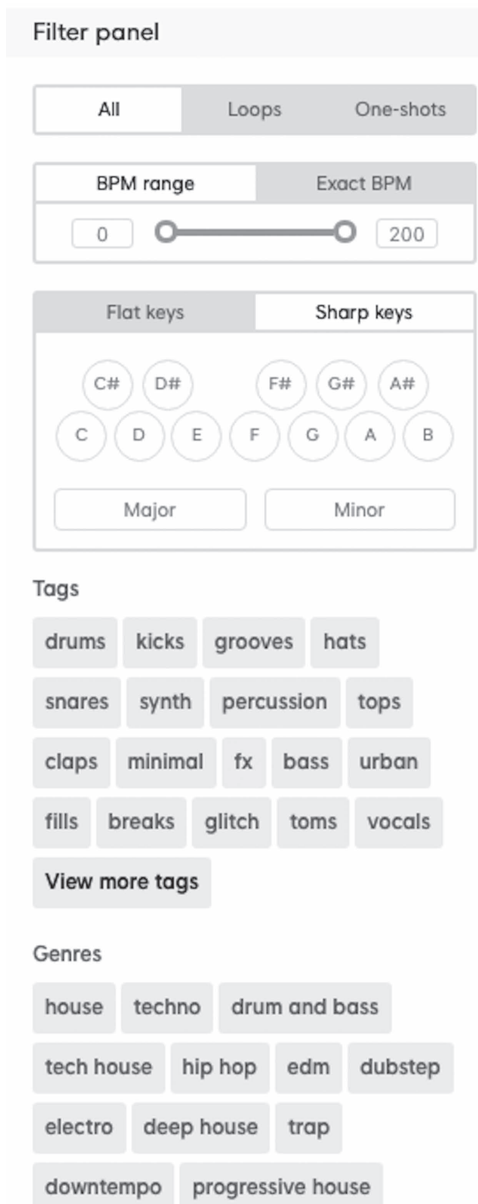


Figure 0.2 Illustrates Splice search filters.

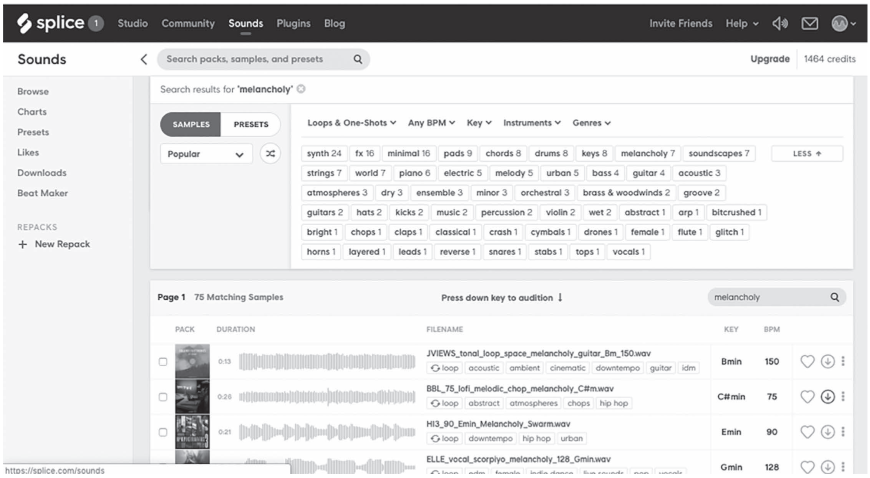


Figure 0.3 Demonstrates the way that Splice handles abstract search terminology.

instance, if I use the search term "kick drum," I will receive results which I can organize by genre, sample length (i.e., a loop or a one-shot sample), sonic description (i.e., distorted, acoustic, synthesized, noise-based, tonal, and others), bpm, genre, and key. Search results can be ordered by popularity, relevance, most recent, or via an entirely *randomized algorithm*. Depending on whether producers are seeking a unique sound, a more well-known sound, or something they perhaps have never heard before, they can select the corresponding option.

The search engine also accepts more abstract terminology (Figure 0.3). For instance, if I search "melancholy" on Splice, it provides 75 sample results (and 11 sample packs).[7] From this point, I can sort through the results by applying various filters. For instance, if I click the filter labelled "pads," the number of search results is reduced to 8. Thus, these CBMP services significantly reduce the amount of time it takes to find the right sounds for any project.

By incorporating these search features into centralized libraries, CBMP services have alleviated a major pain point for music producers. In the past, those who produce beats might sign up for several subscription services through Loopmasters or *Computer Music Magazine*, or they would purchase individual samples or packs through one of many available loop-production services such as Loopmasters, Capsun, and Bingoshakerz. These sample packs were costly, and, in addition, they required producers to constantly sift through numerous web-based sources. While this digitally mimics the tradition of *crate-digging* in vintage hip-hop production—or the act of searching through old vinyl releases for one-shots and loops—it is also an expensive and time-consuming activity for producers. CBMP services greatly reduce the cost associated with purchasing royalty-free samples, as well as the time it takes to find them.

0.4 PROJECT BACKUP AND SHARING

In recent years, it has become a standard practice for studios to keep many terabytes of extra hard drive space available for creating backup copies and to safely store old projects. However, some CBMP platforms have disrupted how producers store these contingency copies of their work. This alleviates a significant stress which producers routinely face, and, as a result, most producers welcome the chance to remove the threat of losing their work.

Some CBMP services offer free and unlimited automatic backup services. Splice, for example, provides the following explanation of their *Splice-sync* feature as follows:[8]

> There are two options for how the app syncs your work, which you can change in the app's preferences page:
>
> 1. (default option) Sync projects in the Splice folder: After installing, a Splice folder and shortcut will automatically be created. Save a project into the Splice folder and the app will automatically sync it. You can also drag and drop existing projects into your Splice Folder to sync them.
> 2. * Sync projects anywhere: If you save a project anywhere on your computer or external drive, Splice will ask if you want to keep it synced with the cloud. You can change the sync setting for a project in the app preferences screen.

Once this backup system is in place, Ableton, Logic, Garageband, and Fruity Loops project files will be automatically uploaded to a remote server every time the user saves their work (Figure 0.4). Alternatively, Loopcloud allows producers to back up their library of audio samples so

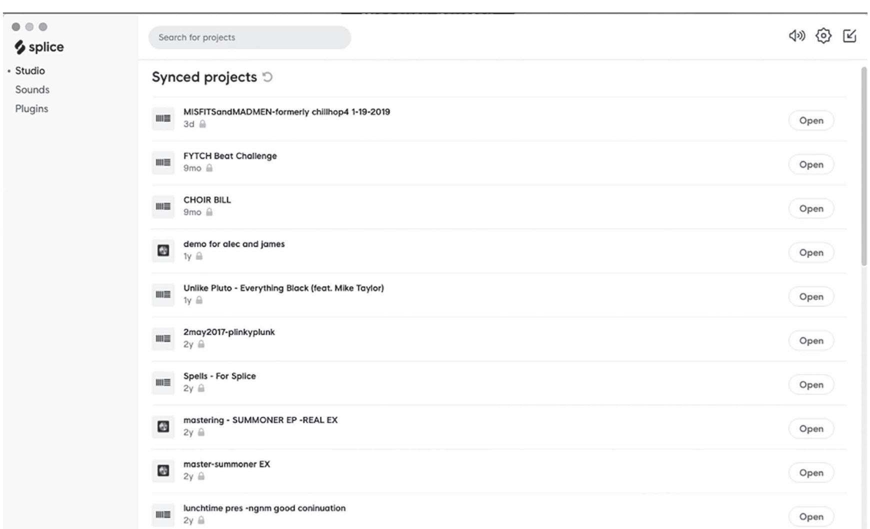


Figure 0.4 Demonstrates the Splice-Sync backup feature in Splice.

that they never have to worry that their creative materials—namely audio samples—will be lost due to a failing computer.

0.5 ACCESS TO PROFESSIONAL SOUNDS

CBMP services make professionally created sounds available to users for a low monthly fee. By using these services, even beginner music producers can familiarize themselves with the types of sonic resources heard on records they listen to. This considerably reduces the learning curve for those seeking to develop production competencies such as sample manipulation, synthesizer programming, editing, and mixing audio. In fact, users can access Just Blaze's drum samples (Jay-Z, Young Guru, Kanye West, Saigon, Eminem, The Game) which are featured on numerous chart-topping hits between 2000–2005 in a kit known as *Meow the Drums* (available on Splice). These samples were taken directly from his AKAI MPC, according to Blaze. In addition, producers such as Zaytoven (Gucci Mane, Migos, Travis Scott), Lex Luger (Gucci Mane, Wakka Flakka, Pusha T, Rick Ross, Juicy J), Getter (OWSLA, Shred Collective, Monstercat), Ducko McFli (Sremmurdrae), and FKi (Post Malone) have also released sample packs on CBMP platforms (Figure 0.5).

While some sample packs are created by famous producers, others are created by lesser-known (or unknown) producers and sound designers. These music professionals, who comprise the majority of contributions to CBMP libraries, create sample packs and synthesizer presets which capture the sonic zeitgeist of existing music scenes (such as hip-hop, trap, house, tech-house, techno, and others). Some of their contributions overtly emulate the sounds of well-known artists. For instance, one sample called "Bpm86_G#_36Chambers_EthnicPluck.wav" clearly draws inspiration from Wu-Tang Clan's cherished album from 1993, *36 Chambers*.

I encourage readers to search for some of the samples from Table 0.1, which draw inspiration from famous producers.

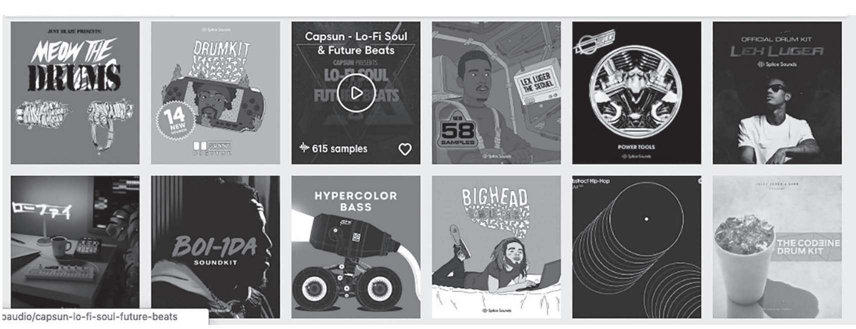


Figure 0.5 Shows some of the sample packs made available by famous musicians such as Just Blaze, Sonny Digital, Lex Luger, Boi-1DA, and Big Head.

Table 0.1 CBMP Samples Inspired by Famous Producers

Influential Producer	*Sample Name*	*Type of Sound*	*Sample Pack*
J Dilla	Bpm85_G_DillaDreamz_Drums03.wav	Drum Loop (continuous)	Shaolin Beats by Prime Loops
J Dilla	cw2_drm90_dillafly_full.wav	Drum Loop (continuous)	Chillwave 2 by Sample Magic
J Dilla	DKVPZ_clap_one_shot_dilla.wav	Snare Sample (one-shot)	Culture Alchemy by Rare Percussion
SwuM	SWUM_percussion_loop_03_150.wav	Drum Loop (continuous)	SwuM Drumkit by Splice (created by SwuM)
Wu Tang/RZA	Bpm86_G#_36Chambers_Snare.wav	Snare Sample (one-shot)	Shaolin Beats by Prime Loops
Just Blaze	JUST_kick_doctor.wav	Kick Sample (one-shot)	Just Blaze—Meow the Drums by Splice (created by Just Blaze)
Lex Luger	LEX_LUGER_clap_trap_slap.wav	Clap Sample (one-shot)	Lex Luger—the Sequel by Splice (created by Lex Luger)
Ducko McFli	Ducko_percussion_loop_boujee.wav	Percussion Loop (continuous)	Ducko McFli Drum Essentials
Boi-1da	BOI1DA_808_08_C	Roland 808-inspired Sub Bass (one-shot)	Boi-1da Soundkit: Bare Sounds for Your Headtop by Splice (created by Boi-1da)
Sonny Digital	SONNY_D_hihat_closed_10.wav	Hi Hat Sample (one-shot)	Sonny Digital Drumkit

0.6 FREE TOOLS/RENT-TO-OWN

Free and inexpensive tools draw many users to CBMP platforms. At the time of writing, for instance, Splice offers users 82 instrument and effects plugins as a membership perk. These plugins come from companies including Xfer Records, iZotope, Valhalla, Voxengo, Plugin Alliance, and many others.

The types of free plugins Splice offers varies widely in selection, from signal processing effects to meters and synthesizers. Similar to samples and presets, free plugins are organized in a searchable database, and users can refine their search results using any of the following *tags*: utility, analog, imaging, modulators, delay, distortion, dynamic, filters, vintage, algorithmic, chorus, EQ, mid side, reverb, stereo field, synthesizer, amp simulator, bit-crusher, mastering, metering. If a producer wanted to find a free vu meter plugin to use on the master bus, for instance, s/he could type in "VU" and filter the results using the search tags "master" and/or "meter."

An advantage to this type of database is that it also allows users to search using tags, even in the absence of user-inputted text. For instance, if someone knows s/he wants the organ sample in a song to sound like it came from a vinyl record but s/he doesn't know exactly what type of plugin to use, s/he can simply click on the search tag for "vintage" or "analog." From this point, one only needs to try the different plugins until the satisfactory sound is achieved.

Another feature of some CBMP services is that they allow users to *rent-to-own* a variety of expensive DAW plugins. Splice, for example, allows users to lease popular synthesizers such as Serum, as well as mixing and mastering VSTs such as Ozone 8 and Neutron 2 ($249). As a result, producers can use these plugins for a small fee ($9.99/month) until the price of the plugin has been paid, at which point payments cease and the user *owns* his/her own copy. Splice plans to grow their plugin rental service in the future (Deahl 2017).

iZotope's Ozone 8, for instance, has a retail price of $499, which may prohibit numerous amateur producers from purchasing the software. However, with more affordable monthly payments, these producers can access iZotope's numerous mastering capabilities, including applications for spatial manipulation, tape emulation, multiband compression, linear phase EQ, and numerous others.

Xfer Records' Serum ($189) is an ultra-high-quality wavetable synthesizer, with an innovative animated graphic interface which allows users to modify wavetables in real time. According to the creators, they have gone to great lengths to ensure that the software is optimized using SSE2 instructions—the standard processing architecture for current Intel chips—in order to provide sophisticated, yet clean-sounding, oscillators.

Upon its release, Serum quickly became a popular plugin with tastemakers and musicians. Both *Electronic Musician* (2014) and *Computer Music* (2015) gave Serum editor's choice awards when it was released, for example. The plugin is also used by numerous famous producers, such as Kanye West, who was caught using a pirated version of the software in 2016 (The Guardian [online] 2016).

0.7 COMMUNITY, CONTESTS, AND COLLABORATION

CBMP services allow digital producers to share their projects with other users, find collaborators, and participate in contests. Users can upload DAW project files from Ableton, Logic, Fruity Loops, and Garageband into a searchable database, and anyone with a Splice membership can access these files. In addition, each database entry specifies the uploader's chosen plugins so that others can determine whether or not they have the corresponding applications. If a user does not have the plugins required for a project, s/he can either purchase (or lease) them through the links provided, or s/he may choose a substitute plugin. S/he can also use the Splice service to message the user and ask for an audio recording of the track in question (rather than MIDI, which is a set of instructions read by a plugin describing pitch, timing, velocity, and related parameters).

These file-sharing capabilities significantly reduce the strain of searching for musical collaborators, which can be a difficult task for not only hip-hop producers but musicians of all types. CBMP community databases allow users connect with one another based on genre type, bpm speed (which is a significant component of many digitally produced genres), DAW type, and a host of custom hashtags. And, users can easily preview projects in the database, making it easier to sort through them (Figure 0.6).

In today's music-making scene, CBMP communities provide a digital space for jamming, sharing, and releasing music. Interestingly, there are no limitations on the types of digital sounds, or music genres, heard within these communities. Although classical music receives little representation on CBMP services, for instance, synthesizer presets and audio samples which reference the genre are prevalent. In particular, CBMP services offer rich possibilities for those who create film, tv, or video game scores.[9]

For many digital music producers, remix competitions provide a chance to gain exposure. Quite often, remix competitions are hosted by established artists or labels for the purpose of promoting their work to other producers, DJs, and tastemakers. While dedicated websites have promoted these events for quite some time, CBMP services also facilitate remix competitions through the provision of tools for community engagement.

For instance, Splice hosts numerous mixing and remixing contests. Some of these contests involve creating new music from a sample pack, such as Brenmar's competition in the Summer of 2017, and some are remixing contests, such as Monstercat's November 2017 competition. For Brenmar's competition, the top four submissions won the chance to collaborate with Brenmar on a cloud-enabled collaborative EP called *The Collab Collection* (2017), the first of its type, which debuted on his record label known as High End Times. Monstercat's competition, on the other hand, offered producers the chance to release a song through their label, as well as a free Roland DJ-808.

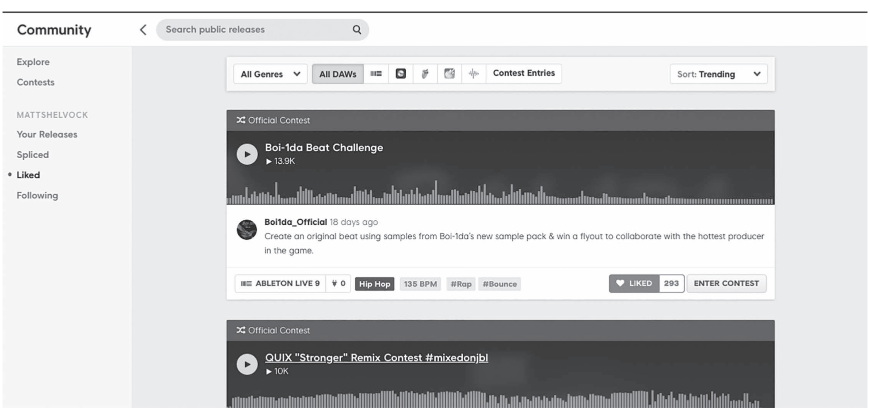


Figure 0.6 Demonstrates the "community" feature on Splice. In the photo above, readers can see some of my "liked" projects, which include a BOI-1DA beat challenge and a remix challenge from QUIX.

More recently, in spring of 2018, a multi-genre producer known as Fytch (bitbird, mau5trap, Dead Beats, Heroic) hosted a beat-making competition which invited users to make music using a folder of samples which he selected in advance. The winner and runner-up both received a private coaching session with Fytch on any topic related to music or music production.

These competitions provide producers with a chance to showcase their work within a larger music community. In many cases, producers who enter these competitions begin to follow one another's music and collaborate with each other. This provides value to these creatives since local amateur concert scenes rarely center around the creation of beats.

0.8 CHAPTER OUTLINE

Now that I have introduced readers to CBMP in a broad sense, Chapters 1–4 will provide a more granular level of information and analysis. In Chapter 1, I provide an overview of sampling and synthesis techniques, as well as an overview of the implications of connecting them with CBMP functionalities. In Chapter 2, I discuss mixing, which is another crucial aspect of beat-making supported by CBMP platforms, before turning the reader's attention towards the creative implications of using CBMP-based *royalty-free* samples (as they are commonly referred to) in music-making in Chapter 3. In this chapter, I provide a detailed explanation and analysis of how digital samples, and sample libraries, are used to produce hip-hop beats—a topic which is rarely explored in print media (if ever). This chapter will be the most useful for readers seeking to understand how samples are used to create new music.

In Chapter 4, after providing readers with some foundational background on mixing practices in Chapter 3, I explore the psychoacoustic dimensions of beat-making. I do so by elucidating some of the cognitive and psychophysical conditions which must be present in order for listeners to *believe in* a record, and I proceed to clarify how beat makers satisfy these conditions. In fact, CBMP services offer audio samples, presets, and mixing tools, which are all created for the purpose of aiding producers in establishing this *believability* on records. In Chapter 4, I explore how producers blend these sonic components to create records, which are, in essence, nothing more than sonic illusions.

The final chapter of this book covers a number of case studies in beat production. These case studies cover popular hip-hop styles including 1990s East Coast/boom bap, G-funk, Dirty South, lo-fi, trap, and experimental. Readers should know that, practically speaking, today's producers blend freely from these styles, as well as others. However, I believe that these categories are sufficient for demonstration, and it is my hope that these case studies inspire readers to engage in creative experimentation with sounds from these sub genres. It is my goal that these case studies will provide a basis from which other creative researchers can build upon.

NOTES

1. www.forbes.com/pictures/eeel45jfeg/1-tiesto-22-million/#6beb644e2cef
2. On many commercially successful hip-hop releases, a single executive producer (either an outside producer or the artist him/herself) will oversee the completion of the record. This individual is often responsible for facilitating recording sessions with emcees, but other responsibilities may include acquiring *beats* by other producers, hiring the required engineers/assistant personnel, or even performing engineering duties.
3. Please see https://splice.com/sounds/labels
4. Moreover, when he does take samples from records, he proceeds to "grab something, take a loop out of it, then flip it, reverse it, distort it, or chop it up. Then I ask myself 'Did I create something?'" (quoted in Pensado's Place 2014). In other words, it would be nearly impossible to identify the samples that Noah 40 uses to produce many hit rap, hip-hop, and R&B songs. When the samples are more easily identified, he simply opts to give rightsholders a portion of his publishing and master rights within his new work (while also paying any applicable fees).
5. This sample can be heard at the following web address: www.whosampled.com/sample/494321/Kendrick-Lamar-FEEL.-Loopmasters-COF-125-Am-LaidOut-Underwater/ accessed September 27, 2017.
6. www.bigfishaudio.com/about.html?1;24;1:::::::::::: accessed October 19, 2017.
7. Over the course of a few months, the number of search results for this more abstract term grew from 58 samples across 7 sample packs to 75 samples and 11 sample packs.
8. Taken from https://support.splice.com/hc/en-us/articles/201178607-Uploading-your-projects-to-Splice accessed December 1, 2017.
9. And, of course, hip-hop is a genre that samples audio from soundtrack sources often, such as Busta Rhymes' use of the *Psycho* (1960) theme in "Gimme Some More."

Plugging In

Exploring Innovation in Plugin Design and Utilization

Andrew Bourbon

INTRODUCTION

Audio plugins have become the sound-processing hub for the increasingly dominant in-the-box (ITB) mixing environment. As the digital audio workstation (DAW) has become the primary environment for the task of music mixing, the breadth and complexity of audio processing tools has grown considerably, covering an ever-expanding range of tools to solve specific problems since the introduction of Waves Q10 in 1992 as the first available third-party plugin.

As plugins have developed, designers have focused their development on several key areas. Emulation of classic hardware has been one of these key areas, with developers continuing to find new techniques for the measurement of classic hardware and the re-creation of that hardware in the digital environment. The market hunger for these tools is clear, with manufacturers continuing to develop for this apparently saturated market sector. In addition to the innovations leading to a higher-quality, more faithful emulation of the hardware, there have been further examples of innovation, allowing functionality that would not necessarily have been available to the hardware user, offering greater flexibility in audio processing and also tacking some of the issues inherent in exploring exact emulations of hardware in a DAW environment where gain staging limitations and contemporary delivery levels are significantly different from those faced in the analog domain at the time of the development of the emulated hardware.

While the innovation is clear in the development of the technologies required to create faithful emulations of complex hardware, there is also an argument that emulation lacks innovation in the development of new control paradigms. The question of emulation and innovation will be discusses in more detail, exploring the potential user base for plugins and the apparent hunger for emulated plugins among the music production community.

One recent development in plugin design inspiration has been to look away from the specific hardware, and instead to the human user associated with record production. Companies such as Waves have developed signature series plugins, which have embraced the more traditional emulation approach, and also explored new interfaces for control that embrace the particular sonic signature associated with a particular engineer. These interfaces have moved away from the traditional interfaces associated with hardware, instead often moving towards a simple interface offering direct control of a semantic descriptor associated with a particular engineer. The celebrity engineer element of plugin innovation has also led to the development of a number of endorsed preset libraries for tools, again designed to give the end user access to the sound of their favorite engineer and record.

Innovation in plugin design can also be found in the interfaces presented to the user, with developers providing new ways of visualizing and controlling the parameters offered by a particular process. In some cases, the aim of the interface is to simplify the controls, providing an interface that makes it as easy as possible for the end user to interact with their sound, without necessarily having a complex understanding of the process behind the interface. Other interfaces provide an entirely different experience to the end user, allowing configuration of every element of the process undertaken in microscopic detail, requiring significant prior knowledge of the audio process in order to take full control of the interface provided. Other manufacturers such as Fabfilter create expert areas in their plugins, offering greater levels of control to users who wish to take greater control of their signal processing through contemporary interfaces that offer significant user feedback through visualization of signal processing.

The final area of innovation that this chapter will explore is in the development of tools for non-traditional audio processing. There has been significant technological innovation in processes such as tuning, spectral noise reduction and manipulation and creative processing, with many of these tools offering functionality that has only become available as users have move into the digital domain and processing power and technology have increased. Companies such as Izotope have created a range of innovative tools exploring visualization, mixing, mastering and noise reduction through the use of technologies such as machine learning. These tools are clearly at the height of innovation in music production, offering users new approaches to both traditional production processes and also to tackle problems that would traditionally be incredibly difficult, or indeed impossible, to fix.

CLASSIC HARDWARE EMULATION: THE SEARCH FOR REALISM

When exploring the plugin portfolio of the major plugin manufacturers, there is a prevalence of emulated hardware tools. Exploring the list of dynamics processors that form the catalog for the Universal Audio UAD

DSP platform reveals a list of compressors that include the following processing tools:

API 2500	Precision Bus Compressor
Manley Vari-Mu	Empirical Labs EL8 Distressor
Fairchild Collection	Elysia Mpressor
Teletronix LA-2A Collection	Teletronix LA-3A
1176 Classic Limiter Collection	Vertigo Sound VSC-2
SSL 4000G	Elysia Alpha
Neve 33609	dbx 160
Tube Tech CL1B	Valley People Dyna-mite
Summit Audio TLA-100A	

The list of compressors here represents an exhaustive list of classic and contemporary studio dynamics processors, all of which have been meticulously emulated by Universal Audio and their partners since their inception as a programming company in 1999. Competitors such as Plugin-Alliance, Softube, Slate Digital and Waves all feature similar product lists, with either direct emulations of much of the hardware noted here or tools inspired by these classics but without the official licensing to allow use of the specific model name.

The demand for high-quality emulations is clear, with only one of the plugins named representing an original design, with the Precision Bus Compressor conceived with the specific aim of offering clean gain reduction from a voltage-controlled amplifier (VCA) style compressor. The 1176 is established as a studio standard (Massy, 2016) and is one of the most emulated compressors as a plugin, with all the major manufacturers offering emulations and native versions of these compressors available in most major DAWs, and is famous for its use of a field effect transistor (FET) in the gain control circuit (Case, 2007). The 1176 is a particularly colorful compressor, offering very fast attack times, with a real sense of "hair" added to the sound when in compression and a sense of "air" provided as the release speed is increased. Commonly used for vocal processing, bass processing when harmonic enhancement is desirable, and also on drums, Universal Audio has revisited the 1176 in their 1176 limiter collection, adding new improved emulations to the already existing legacy version. The main focus of the upgraded emulation has been to model the complete circuit path of the compressor, including the transformers and amplifiers in the input and output circuitry. Through measurement and subjective auditioning of the compressor, the increase in harmonic complexity is clear, with resultant increases in harmonic content both in compression and out of compression, confirming the enhanced impact of the complete modeled circuits on the sound. Universal Audio has also provided three different editions of the 1176, with a Rev A, Rev E and AE versions of the compressor available to the end user having been through the component modeling process.

The three modeled versions provide three different sets of sonic characteristics, with each having different impact on the sounds that they are processing. The Rev A, for example, exhibits higher distortion, and also has a different release profile from the Rev E due to the variation in the program-dependent release behavior.

Recent innovation in creating realistic emulation of hardware performance has seen even more detailed modeling taking place. The world-famous SSL 4000 mixing console has been modeled by a number of manufacturers, with companies such as Waves, Universal Audio and Softube modeling the channel drive either as part of the channel strip or as additional components, allowing the user to build a full SSL channel strip in the digital environment. Massey describes the process of "creaming the mic pres", in order to bring excitement into recordings through the dynamic compression and distortion associated with various vintage circuits. Universal Audio have also modeled the performance of the dbx voltage VCA found in the channel output section, as well as adding the optional Jensen transformer on the microphone input, offering the user the full SSL channel experience within the DAW with full control over gain staging, as found in the hardware and the tonal options to explore the classic techniques associated with creative abuse of the channel strip.

Brainworx have been responsible for further innovation in their emulations of contemporary channel strips. Through their process of Tolerance Modeling Technology (TMT), Brainworx have measured the component tolerances in every piece of the channel strip, and then realized those variations in the console, with significant and subtle variations present throughout the 72 TMT channels. This process has been competed for three different consoles, with a Neve VXS and SSL E and G series all measured and emulated. Control is provided to the user, allowing either random allocation of channels or specific choices to be made, with phase variation and stated value changes in the channel strip offering changes to the presentation of audio, particularly when applied to stereo material. The impact of this randomization can be significant, resulting in perceived changes often considered desirable by end users, with increased image "width", "depth" and "punch" experienced by the listener.

Component circuit modeling has been a significant innovation in the recent history of plugin design. The TMT technology and component modeling used by other manufacturers has seen a significant increase in the performance of emulations. Digital audio consoles primarily designed for live sound such as the Midas Pro series have plugins built into the console developed from modeling their own hardware, built specifically for the purpose of digital modeling, and later released by Klark Technik as the Square One Dynamics processor in order to offer the best possible real-time performance to engineers. Waves and Universal Audio are also incorporating their plugins into dedicated digital signal processing (DSP) units to be integrated into the digital console workflow with a minimum of latency. Improvements in reinforcement systems are seeing users focusing further on the signal processing engaged in live mixing, turning to the emulated processors commonly used in the studio environment to re-create the signature sounds from records in real-time live performance.

It is important to note the influence of convolution technology on plugin design, and in particular on the development of reverb processing in the DAW. Products such as Altiverb offer engineers an extensive collection of classic reverb units and real spaces, all captured through convolution technology. Universal Audio is using combinations of algorithmic delay networks and convolution in their BX spring reverb in order to faithfully re-create the unique sound of the AKG spring.

Perhaps the most innovative use of a convolution approach is being pioneered by Acustica audio, who are using dynamic convolution to create emulations of a range of classic processing tools. Though the DSP load of this technology on the host machine is significant, and there are some associated complications in the real-time manipulation of parameters, the sonic quality afforded by this technology is impressive and represents an exciting innovation in emulation methodology.

REALITY RULES: SKEUOMORPHISM, NOSTALGIA AND PLACE

As the DAW has become the primary interface for the creation and mixing of music, there is now a generation of engineers who, rather than moving from the analog studio into the digital domain, have their entire studio experience in a computer-based environment. Despite not having had access to hardware tools, many young engineers do have an appreciation of classic hardware tools as explored in the digital domain. As already discussed, much of the innovation in plugin emulation has been focused on creating ever more accurate models of hardware. In the digital domain, the changes in the process of capturing, processing and generating sound have led to challenges for those exploring these emulated tools without the experience of using these tools in an analog domain. It is possible when working in the modern DAW to almost completely ignore traditional gain staging approaches, driving channels beyond the levels achievable using vintage recording equipment and simply turning down the master fader to avoid output clipping. Floating-point processing has led to internal clipping simply not being an issue; as long as the signal at the output is not clipping the digital-to-analog converter, the signal will remain clear of distortion. As music has become louder at distribution, young engineers are creating pre-mastered mixes that are already at levels above those that would have been traditionally provided pre-mastering. Many of the emulated hardware tools will react aggressively to this signal level, with tools such as Fairchild compressors, for example, in significant gain reduction with the threshold set at a high level and the input gain low; manufacturer-driven solutions to this contemporary gain staging issue will be discussed later in this chapter.

The proliferation of tape processing tools, classic console processing and vintage hardware in the digital domain has also influenced the sense of nostalgia for the tools in the hardware environment, and in turn is feeding the constant increase in emulated tools in the DAW. This is particularly interesting with engineers who have only experienced music production in a digital environment, yet still feel a sense of nostalgia towards analog technology that there is no physical experience of using. The incorporation of arguably undesirable elements such as noise have added to the digital emulation experience. Bell et al. (2015) state in regards to Slate VTM: "The VTM makes digital *look* vintage, but more importantly, it makes digital *sound* vintage, going so far as to offer optional tape hiss. This is 'technostalgia' at its apogee, and it extends our definition of skeuomorphism beyond visual to aural". The impact of tech nostalgia has been further enhanced by the significant rebirth of vinyl as a format, with new challenge faced by those cutting vinyl from the aggressive, bright and loud masters generated in contemporary popular music. The integration of classic processing in the contemporary DAW environment provides a significant challenge to educators and enthusiasts looking to embrace these tools without the limitations enforced by the analog environment.

The influence of nostalgia on plugin design innovation is not limited to the emulation of classic tools, but also with the places in which those tools could be found. Waves, for example, offer a series of *Abbey Road*–inspired plugins, Universal Audio offer access to an emulation of the Ocean Way recording rooms, and Eventide offer Tverb, providing the Hansa studio reverb made famous during the recording of David Bowie's *Heroes*.

All of the emulated places and hardware tools have also seen significant development in the look of the plugin, embracing literal re-creations of the hardware visually. This skeuomorphic approach provides the user with an interface that affords the same control as the real hardware, with the quality of the visual impact of the interface an important part of the user experience. Slate Digital plugins, for example, exhibit different levels of rack rash and general damage each time a plugin is instantiated. Though this clearly has no effect on the sound of the plugin, it is clearly important to the end user and as such represents an important innovation in plugin design.

ALTERNATIVE REALITY: ADDING FUNCTIONALITY IN EMULATIONS

Though the focus in emulation innovation has been in creating ever more accurate emulations of hardware, manufacturers have also been adding functionality beyond that found in the original units. The tolerance modeling technology found in the Brainworx console plugins, for example, can be varied randomly, providing variations not originally enacted by the user. The impact of this random processing is a change in image size and depth, creating according to Brainworx, a "hyper realistic mixed on a big console sound for the first time in the box". There is also user control over other features, including a total harmonic distortion (THD) control adding variable color, configurable per track. Wet dry functionality is added to the dynamics, as well as a secondary release time to customize the release of the compressor and avoid audible pumping. Middle and side monitoring functionality is provided in the plugin, with additional gain staging control to that of the original console to manage noise floor and color through the plugin itself. This concept of hyper-reality, where the user gets a level of control beyond the hardware original is a key innovation, both in the plugin market and also in the hardware market. The SSL G Bus Compressor is available as a DIY project and has also been created by a number of other manufacturers, often adding high pass filtering in the sidechain to avoid the compressor reacting to low end and compromising the lower octaves that have become so important in contemporary popular music production. Similarly, the option to modify the sidechain stereo linking is provided in a number of hardware tools, with the ability to use a summed sidechain or true stereo detection, resulting in either perceived width or central punch, depending on the mode selected. Plugin compressors are innovating with even more sidechain modification potential, which will be explored later in this chapter when exploring contemporary non-emulated interfaces.

Another recent innovation in plugin design has seen the ability for the end user to modify the calibration of the plugin itself. The levels of signals inside the DAW can be incredibly variable, depending on engineer approach, genre and positioning of the audio processing in the audio chain. By default, the Fairchild 660 plugin is calibrated so that 0 dBFS will result in what would be approximately +20 dBu into the hardware unit. At this input level, there would be significant coloration in the audio circuit, and the unit would be subject to high levels of gain reduction. In order to control this, it is possible to modify the reference level from the default 16 dB, with a range of variation in calibration from 4 to 28 dB. This creates a range of control, allowing subtle coloration and control over small amounts of compression that is useful in bus processing and mastering to more aggressive channel and parallel bus usage. It is also of note that Universal Audio have added wet dry control and sidechain filtering to the parameters traditionally offered in a Fairchild compressor, providing yet more innovative control in addition to the default accurate emulation.

Tape processing has become more prevalent in emulation, with significant steps being made by manufacturers in the emulation of magnetic performance. The range of control offered in tape plugins varies significantly, depending on the provenance of the tape tool being employed. Universal Audio, for example, have created faithful emulations of the Ampex ATR-102 and Studer A800 tape machines, and have also created a simplified tape machine with a focus on ease of use and for real-time processing at the point of capture using their own DSP interfaces. The two emulated tape machines both offer full calibration control, with the ability to modify noise, sync and repro equalization and level. The manual provides the recommended calibration for the four tape formulations provided within the emulation. It is possible to vary these parameters, allowing a level of flexibility in tape processing that would not have been viable to instantly compare on an analog machine. It is important to note that the headroom control found on the latest dynamic processors is not available in this tape emulation, meaning that the end user has to be very aware of the gain staging within their production depending on how the tape is to be driven and where the tape sits in the processing chain. As has already been discussed in this chapter, there is a generation of young engineers who have only experienced audio processing in a digital environment, and who have knowledge of classic processing gained only through the use of emulation tools. With emulations such as the Studer A800, a knowledge of gain staging related to traditional studio practice is desirable in order to maximize the potential of the tape in the classical sense. It is also true to say that one of the great innovations in contemporary production comes from what would be seen in traditional approaches as a creative misuse (Keep, 2005) of tools such as tape, where traditional approaches would not afford the processing opportunities offered in the DAW.

As well as specific emulated tape machines, manufacturers such as U-He have provided an interface that appears to the user as a tape machine, but offers control over a greater range of parameters without reference to classic machines and tape formulas. Two formulas are offered, simply described as vintage and modern. Tape speed control is provided, but unusually is continuously variable to allow specific dialing in of tape response. A pre-emphasis control is also provided and offers control over transient response and tone, and a compander section allows exploration of noise reduction technology as used and misused in the traditional analog studio both to add dynamic range and famously in the processing of backing vocals to create a sense of "air" and "height" using Dolby A processing. Further innovation is provided through control of parameters of tape processing often considered to have negative connotations. It is possible to manipulate noise performance, wow and flutter and a range of advanced features such as asperity, crosstalk and bias. It is also possible to manipulate the emulated physical characteristics of the tape machine. The captured frequency response is significantly impacted by the size of the reproduction head gap, with this physical characteristic fixed in an analog tape machine. The head gap width on a mastering machine, for example, will tend to be very small to allow high-frequency extension. Companies such as U-He allow manipulation of this physical feature, along with other characteristics such as bias and the tape bump, giving the user exceptional levels of control over their audio signal both in terms of frequency response and distortion characteristics. Through manipulating these parameters, users are able to re-create their desired tape machine characteristics to facilitate more traditional tape processing with enhanced control, and also engage in creative processing not easily achieved with an analog machine.

PICK-AND-MIX PROCESSING: BUILD YOUR OWN TONE

Tape processing provides clear examples of manufacturers creating a pick-and-mix approach to plugin design, with users able to combine elements of different emulated units into a single signal path. Similar functionality can be found in EQ units, where the bands can each feature different EQ curves to represent different emulations in different frequency bands. It is possible in a single interface to select a Neve filter, with mid-range bands from an SSL and the top end from an APL 550, for example. This process has been taken further by Waves, who as part of their Manny Marroquin signature series provide an EQ that features the preferred hardware EQ choices for specific frequency ranges as employed by Manny Marroquin. The influence of mix engineers on plugin design will be discussed in more detail in the following section.

As well as creating emulations that are faithful to the original machines, it is clear that extreme control over traditional processes sits at the heart of plugin innovation and use in contemporary processing, with the majority of modifications made to support use in a contemporary production environment. Soundtoys, for example, use generic interfaces that contain processing elements inspired by hardware. Decapitator, for example, is a distortion plugin that offers multiple distortion circuits from devices, including the germanium stage of a Neve preamp and the tube distortion found in the Thermionic Culture Culture Vulture. Drive and EQ controls allow the characteristics of the distortion to be controlled, with an additional "punish" control driving the distortion to relatively extreme levels. The high-order distortion created by this approach has become more prevalent in contemporary production as creative abuse has seen high levels of distortion produced and captured with high fidelity. There are also a number of commercially available products such as Dada Life The Sausage Fattener embracing the heavily compressed and distorted aesthetic associated with particular musical genres and specific artists with a greatly reduced set of interface controls designed for instant musical gratification.

MIX ENGINEER INSPIRED PLUGIN DESIGN: SEMANTIC DESCRIPTORS

As well as being inspired by classic hardware tools, plugin manufacturers such as Waves have also embraced contemporary mix engineers in their plugin design approach. The Waves signature series tools embrace a range of engineers, including Chris Lord Alge (CLA), Manny Marroquin, Andrew Scheps, Jack Joseph Puig (JJP), Eddie Kramer and numerous other engineers associated with contemporary pop and rock music. These signature series tools in combination with online resources such as Mix with the Masters, Slate Audio Legends and Pensado's Place have all led to mix and recording engineers becoming recognized names, and their sonic signatures becoming known by engineers, producers, musicians and consumers.

The CLA Classic Compressors and JJP Analogue Legends are examples of Waves signature plugin bundles that focus on emulation of classic hardware associated with the engineers, and have continued the tradition of hardware emulation already discussed in this chapter, only this time bringing the sonic personality of the engineer into consideration. Perhaps a more interesting innovation has taken place in the creation of the signature series plugins, which rather than providing the tools found in the studios of the engineers focus more on the sonic impact of the processing undertaken by these engineers.

The Waves CLA signature tools target specific elements within the mix through a series of faders, each focusing on a particular characteristic of the target sound; it has controls for bass, treble, compression, reverb, delay and pitch. Each of these faders has three options for enhancement, with the use of some semantic descriptors for each of the processes. The treble control, for example, offers options such as "bite", "top" and "roof" rather than citing specific frequencies or tools that would be used by CLA to add a sense of "bite" or "roof" to a vocal. The compress fader offers options such as "push", "spank" and "wall", all of which conjure strong images of how the sound may well be manipulated by the fader. This use of semantic descriptor is an interesting development in plugin design, and is certainly something that could be expected to develop as manufacturers run out of tools to emulate and are forced to find new methods of innovation in order to maintain the commercial success of their plugins.

The Waves CLA Drums module interface is very similar to the Vocals interface, with the most significant difference being the choice of drums to be processed using the dial located on the left of the plugin. The other significant difference is the removal of the pitch stereo widening tool in favor of a gate, which has a simple hard or soft setting with a fader to control the impact of the gate on the target sound. Very similar descriptors are used, with some variations in the reverb choices, again to represent the process undertaken by CLA in his own mix approach, as documented in detail in both Mix with the Masters (*Mix With The Masters*, 2018) and Slate Audio Legends (*Audio Legends – Slate Digital*, 2018) resources.

The chosen descriptors found in these tools very much align with the sonic signatures of a CLA mix, providing a user with direct tools that bring the immediacy and impact of a CLA mix to their own production, or indeed the impact of the various engineers who have signature processing available. The ability to mix and match different signature plugins also has exciting potential in utilization, with engineers able to use the audio fingerprint or 'auroprint' as an influence for an engineer finding their own sound through semantic descriptors rather than targeting a single engineer tone or set of classic emulated hardware tools.

Within the signature series of plugins, there are also tools that explore alternative approaches to audio processing. As is suggested in the name of the Waves 'Scheps Parallel Particles' plugin, this tool offers a set of present parallel processing approaches that are designed to add thickness, air, bite and sub, with only the sub processing offering control beyond a simple send. The Andrew Scheps mix approach is discussed in depth in Mix with the Masters (*Mix With The Masters*, 2018), with his mix template available to registered users. This template shows the reliance on parallel processing in his mix system, and this plugin provides a tool that provides the desired processing results in a simplified interface, again driven by semantic descriptors without requirement for knowledge of the process, only the imagination to use the tools to create the sound desired by the end user.

The Manny Marroquin bundle is the most traditional of the processing bundles available from the Waves Signature Series. The EQ, for example, provides a set of EQ filters that represent the tools regularly used by Manny Marroquin in his mix process. The reverb captures settings from the reverbs used in his sessions, and the delay is a relatively common setup but with some additional processing options including distortion and reverb within the delay plugin itself. The Marroquin Signature Series very much represents the mix approach taken by Manny Marroquin, affording a great deal of user control while still embracing the associated processing approach.

Semantic descriptors have become an important feature in the analysis of mixes, leading to the development of tools for automatic mixing that use instrument tags and semantic descriptors to drive the automated mixing process (De Man and Reiss, 2013). Though not the focus of this chapter, feature extraction and knowledge-engineered automatic mixing solutions are becoming increasingly prevalent in plugin design. Equalization, panning and dynamic processing are all undertaken by automatic mixing tools, with companies such as Sound Radix offering tools for automatic adjustment of phase of elements within a mix that could not be achieved without the algorithmic analysis undertaken by the plugin.

It is also noteworthy that there has been a proliferation of preset libraries provided by engineers for more traditional processing tools. As well as providing an effective endorsement for the processing plugin in question, these presets also provide a useful starting point for both experienced and novice engineers. The market demand for presets is also a significant contributor to the proliferation of preset lists, with the reliance on presets in synthesis, for example, well established through the development of complex synthesis technology in the 1980s (Théberge, 1997) and still present in today's plugins. As discussed in this chapter, one of the challenges faced by plugin developers is around issues of gain staging and headroom, with presets often likely to be influenced by the gain staging of the user themselves, and as such will still potentially require modification in order to achieve the desired outcome.

CONTEMPORARY INTERFACES

The Waves Signature Series plugins represent a new approach to plugin innovation, embracing a simplified approach to processing driven by signature sounds of engineers and via an interface that has taken a step away from traditional hardware interfaces. Plugin manufacturers such as Fulfiller and Soundtoys have taken the decision to move away entirely from the traditional interfaces associated with hardware, and instead have focused on creating digital tools specifically to provide the maximum control and feedback to the user in the DAW environment. DAWs have offered a range of native processing, with the traditional DAW EQ and dynamics strip representative of the typical digital channel strip also found in the majority of digital consoles. In the case of dynamics, the classic digital channel strip demonstrates a great deal of information, showing the compression curve, giving clear indication as to the status of the compression and providing easy access to parameters such as ratio, attack and release. In contrast, the basic dynamics - processing strip found in Pro Tools offers very little in the way of tonal control.

The Fabfilter Pro C-2 interface, which although has some similarities with the traditional digital processing found natively in that DAW, offers significantly more detailed control and feedback. As well as offering tonal choices via the semantic descriptors found in the mode control, Pro C-2 provides detailed metering, giving visual feedback on the action of the compressor. The knee of the compressor can be manipulated using a high-resolution slider, the gain reduction range can be limited, and wet dry controls are provided. There are also a series of advanced control options in the expert area, allowing sidechain detection linking control, and multi-point filtering of the sidechain signal, offering the end user significant control over the action of the compressor and the impact on the resulting audio. It is also interesting to analyze the range of preset options available in this compressor, and indeed the potential for users to study the settings that are used to create these presets in further enhancing their ability to create their own approaches.

DMG Audio Compassion is another tool that has been developed with the aim of putting full control of the audio processing in the hands of the end user, offering a wide variety of control parameters that require an in-depth understanding of compression and compressors. The manual describes every parameter in detail, and provides insight into compression for even the most experienced engineers. DMG Equilibrium is similar to Compassion in that it places its focus in providing full user control, with the interface setup dictated by a configuration wizard when instantiating the plugin. Fabfilter and DMG Audio are examples of two companies who are developing tools that place control and metering of audio signal at the heart of their design, with Fabfilter focusing on a work-friendly user interface and DMG providing maximum control and technical detail. For the experienced engineer who enjoys dialing in their tools, these innovative plugins offer huge processing potential, with the quality of processing on offer matching the quality of the interface.

RESTORATION AND TUNING

Audio restoration and tuning has become increasingly important in the development of contemporary workflows and sounds. Paterson (2017) discusses the use of auto tune in contemporary production, with the process of auto tune being used beyond simple correction and instead becoming a detectable sonic signature among the general populace. There has been significant innovation in the processes of detection and control of tuning, with Celemony Melodyne offering the ability to re-voice chords taken from a single recording, manipulate formats in vocal recordings and to facilitate complex sound design through manipulation of an existing signal.

Izotope have developed a series of DAW plugins and stand-alone applications, designed to support mixing, mastering and spectral restoration and manipulation. The noise reduction control offered by Izotope has become industry leading, and has been developed through techniques of machine learning to identify and manipulate noise, with impact from the removal of unwanted noise on music recordings to the addition of location ambience to ADR in postproduction.

HYBRID APPROACHES AND HARDWARE INTEGRATION

Recent innovation has seen an increase in the development of audio products that provide a hybrid approach to processing. The Universal Audio Apollo interfaces, for example, feature the Unison microphone preamp technology, where the analog characteristics of the preamp work alongside a DSP-based plugin to provide a deeper level of integration and realism in the emulation of classic studio hardware. Townsend Labs have created a hybrid microphone system, which uses a specific microphone in combination with a plugin to model the complex characteristics of classic microphones, including the previously not emulated off-axis response and proximity effect.

Other companies are exploring the benefits of recall by offering plugin-based control and recall of hardware settings. Total recall is identified as one of the core advantages of mixing in the box (Paterson, 2017), with plugin control also offering complex automation opportunity not traditionally offered by hardware recall systems. Tegeler Audio provide a range of hardware tools including reverb, compression, equalization and full channel strips with full DAW integration, with Bettermaker going a step further in offering hardware that can only be controlled via plugin interface.

Hardware control over digital re-creations of classic hardware has provided challenges to manufacturers, with many building on the HUI protocol to create generic tools for the control of the DAW environment. In recent years, Eucon has seen a number of innovations, allowing manufacturers to create mappings for complex control using Eucon-enabled devices for direct parameter control. One of the difficulties faced by users in exploring these hardware control interfaces is in their generic layout, with the hardware failing to emulate the layout that has been so specifically re-created in the plugin interface. Softube have developed an innovative solution to this problem, providing a hardware channel strip that is capable of supporting a range of classic consoles in a layout that offers familiarity and a clear link to the emulated channels. Though workflow integration across multiple platforms remains a challenge, the development of dedicated hardware for the control of digital channel strips marks an interesting and powerful development for engineers for whom consistency of workflow and layout are important factors.

THE FUTURE: INNOVATION, EMULATION AND RE-EDUCATION

The impact of innovation in plugin design since 1982 has played a significant role in defining the DAW as the primary environment for music creation and mixing. There is clear potential for the development of new interfaces, but these rely on developments to take place in the entire control paradigm explored by the DAW. New techniques have seen hardware emulation improve to the point where many professional engineers feel entirely comfortable replicating their traditional mix approach while mixing ITB. Although we have seen further developments, particularly in the creation of digital tools that offer a level of control and processing potential that would previously have been unavailable, a number of questions still need to be answered as to the future developments of plugins.

As the quality of emulation reaches a plateau, and the range of desirable tools available to be emulated reduces, designers are faced with new challenges in satisfying the market demand for improvements and innovation. Despite the developments in semantic-driven processing and innovative interfaces, the majority of focus in the industry still falls on emulation and addition of features to provide flexibility and control around defined sonic signatures. Strachan (2017) discusses concepts of democratization of technology, but the addition of features to classic processing tools is arguably more likely to see a gentrification of such processing, with the required knowledge of not just the specific tool but also the advanced control and calibration of that tool required to successfully replicate the mix approaches established in popular music production practice. With companies such as Slate Digital developing engineer-led education programs, perhaps the next innovation is not in interface design or emulation quality, but in the methodologies explored in educating the market in the use of plugins and their broader utilization in mix and production practice.

BIBLIOGRAPHY

Audio Legends – Slate Digital (2018). Available at: https://slatedigital.zendesk.com/hc/en-us/categories/115001623068-Audio-Legends (Accessed: 1 April 2019).

Bell, A., Hein, E. and Ratcliffe, J. (2015). Beyond skeuomorphism: The evolution of music production software user interface metaphors. *Journal on the Art of Record Production*, 9.

Case, A. (2007). *Sound FX: Unlocking the creative potential of recording studio effects*. New York: Focal Press.

De Man, B. and Reiss, J. (2013). A knowledge-engineered autonomous mixing system. In: *Proceedings of the 135th audio engineering society conference*, New York, 17–20 Oct.

DMG Audio. (n.d.). *Compassion manual*. Available at: https://dmgaudio.com/dl/DMGAudio_Compassion_Manual.pdf [Accessed 14 Jan. 2018].

Fabfilter Software Instruments. (n.d.). *Pro.C2*. Available at: www.fabfilter.com/products/pro-c-2-compressor-plug-in [Accessed 14 Jan. 2018].

Massy, S. (2016). *Recording unhinged*. Milwaukee: Hal Leonard.
Milner, G. (2009). *Perfecting sound forever: An aural history of recorded music*. New York: Faber and Faber.
Mix With The Masters (2018). Available at: https://mixwiththemasters.com/node (Accessed: 1 April 2019).
Paterson, J. (2017). Mixing in the box. In: R. Hepworth Sawyer and J. Hodgson eds., *Mixing music*. New York: Routledge.
Schmidt Horning, S. (2013). *Chasing sound: Technology, culture & the art of studio recording from Edison to the LP*. Baltimore: The John Hopkins University Press.
Strachan, R. (2017). *Sonic technologies. Popular music, digital culture and the creative process*. New York: Bloomsbury Academic.
Théberge, P. (1997). *Any sound you can imagine*. Hannover: Wesleyan University Press.
Universal Audio. (n.d.). *UAD plug-ins*. Available at: www.uaudio.com/uad-plugins.html [Accessed 14 Jan. 2018].
Waves. (n.d.). *Signature series bundles*. Available at: https://dmgaudio.com/gallery_all.php [Accessed 14 Jan. 2018].

Introduction

"Shannon wants to feed not just data to a Brain, but cultural things! He wants to play music to it!"

Alan Turing (father of modern computing) about Claude Shannon (father of information theory), during his 1943 visit to Bell Labs

In this chapter, we give an overview of the history and the state of the art in intelligent audio and music production, with a particular emphasis on the psychoacoustics of mixing multitrack audio, automatic mixing systems and intelligent interfaces.

1.1 Intelligent Music Production – An Emerging Field

Recent years have seen the emergence of intelligent systems aimed at algorithmic approaches to mixing multitrack audio with only minimal intervention by a sound engineer. They use techniques from knowledge engineering, psychoacoustics, perceptual evaluation and machine learning to automate many aspects of the music production process.

An increasing number of companies – from startups to major players in audio software development – have released new products featuring high-level control of audio features, automatic parameter setting, and full 'black-box' music production services. Many of the recent conferences, conventions and workshops by the Audio Engineering Society have featured dedicated sessions on topics like semantic music production or intelligent sound engineering. All of this leaves little doubt about the importance, in both academia and industry, of the wider field of analysis and the automation of music production processes. Existing approaches to problems in these key areas are underdeveloped, and our understanding of underlying systems is limited.

For progress towards intelligent systems in this domain, significant problems must be overcome that have not yet been tackled by the research community. Most state of the art audio signal processing techniques focus on single channel signals. Yet multichannel or multitrack signals are pervasive, and the interaction between channels plays a critical role in audio production quality. This issue has been addressed in the context of audio source separation research, but the challenge in source separation is generally dependent on how the sources were mixed, not on the respective content of each source. Multichannel signal processing techniques are well-established, but they are usually concerned with extracting information about sources from several received signals, and not necessarily about the facilitation or automation of tasks in the audio engineering pipeline, with the intention of developing high-quality audio content.

Thus a new approach is needed. Our goal in most music production tasks relates to the manipulation of the content, not recovering the content. Intelligent Music Production (IMP) has introduced the concept of multitrack signal processing, which is concerned with exploiting the relationships between signals in order to create new content satisfying some objective. Novel, multi-input multi-output audio signal processing methods are required, which can analyze the content of all sources to then improve the quality of capturing, altering and combining multitrack audio.

This field of research has led to the development of innovative new interfaces for production, allowing new paradigms of mixing to arise. However, it has also uncovered many gaps in our knowledge, such as a limited understanding of best practices in music production, and an inability of formal auditory models to predict the more intricate perceptual aspects of a mix. Machine learning has shown great potential for filling in these gaps, or offering alternative approaches. But such methods often have the requirement of being tailored towards problems and practical applications in the domain of audio production.

The following sections present an overview of recent developments in this area.

1.2 Scope

The music lifecycle, from creation to consumption, consists loosely of composition, performance, recording, mixing, mastering, distribution and playback. For the initial music creation stages of the workflow, generative music and algorithmic composition have shown the potential of autonomous music creation systems. Papadopoulos et al. [1] provide an excellent review of twentieth century approaches, and more recently there has been a surge in generative and automatic music composition for everything from elevator music to advertising. It is highly likely that this will become an increasing part of casual music consumption. The latter stages of the workflow, related to music consumption, have already been transformed. On the distribution side, musicians can share their own content at very little cost and effort, and intelligent recommendation systems can find preferred content and create bespoke playlists based on the user's listening history, environment and mood.

1.2.1 Intelligent

By intelligent, we mean that these are expert systems that perceive, reason, learn and act intelligently. This implies that they must analyze the signals upon which they act, dynamically adapt to audio inputs and sound scenes, automatically configure parameter settings, and exploit best practices in sound engineering to modify the signals appropriately. They derive the processing parameters for recordings or live audio based on features extracted from the audio content, and based on objective and perceptual criteria. In parallel, intelligent audio production interfaces have arisen that guide the user, learn their preferences and present intuitive, perceptually relevant controls.

1.2.2 Music

Many of the concepts described herein might also be widely applicable in other audio production tasks. For instance, intelligent mixing technologies could have strong relevance to game audio, where a large number of sound sources need to be played simultaneously and manipulated interactively, and there is no human sound engineer in the games console. Similarly, they are relevant to film sound design, where Foley, dialog and music all need to be mixed, and low budget film and TV productions rarely have the resources to achieve this at a very high standard. However, to keep focus, we assume that the problems and applications are restricted to music.

1.2.3 Production

While there is overlap, we are not specifically referring to music creation – the composition and performance. Nor are we concerned with the distribution and consumption that happens after production. The middle stages of the music workflow – recording, mixing and mastering – are all about the production of the music. They are concerned about how the creative content should be captured, edited and enhanced before distribution.

1.3 Motivation and Justification

The democratization of music technology has allowed musicians to produce music on limited budgets, putting decent results within reach of anyone who has access to a laptop, a microphone, and the abundance of free software on the web [3, 4]. Despite this, a skilled mix engineer is still needed in order to deliver professional-standard material [5].

Raw, recorded tracks almost always require a considerable amount of processing before being ready for distribution, such as balancing, panning, equalization, dynamic range compression, and artificial reverberation to name a few. Furthermore, an amateur musician or inexperienced recording engineer will often cause sonic problems while recording. As noted by Bromham [6], "the typical home studio is entirely unsuitable for mixing records, so there is a greater need than ever to grasp how acoustics will impact our environment and how to work around these inherent shortcomings." Uninformed microphone placement, an unsuitable recording environment, or simply a poor performance or instrument further increases the need for an expert mix engineer [7].

In live situations, especially in small venues, the mixing task is particularly demanding and crucial, due to problems such as acoustic feedback, room resonances and poor equipment. In such cases, having a competent operator at the desk is the exception rather than the rule.

These observations, described in further detail in [114], indicate that there is a clear need for systems that take care of the mixing stage of music production for live and recording situations. By obtaining a high-quality mix quickly and autonomously, home recording becomes more affordable, smaller music venues are freed from the need for expert operators for their front of house and monitor systems, and musicians can increase their productivity and focus on the creative aspects of music production.

Professional audio engineers are often under pressure to produce high-quality content quickly and at low cost [8]. While they may be unlikely to relinquish control entirely to autonomous mix software, assistance with tedious, time-consuming tasks would be highly beneficial. This can be implemented via more powerful, intelligent, responsive, intuitive algorithms and interfaces [9].

Throughout the history of technology, innovation has been met with resistance and skepticism, in particular from professional users who fear seeing their roles disrupted or made obsolete. Music production technology may be especially susceptible to this kind of opposition, as it is characterized by a tendency towards nostalgia [6, 10], and it is concerned with aesthetic value in addition to technical excellence and efficiency. The introduction of artificial intelligence into the field of audio engineering has prompted an outcry from practitioners who reject the concept of previously manual components of their jobs being automated [11].

However, the evolution of music is intrinsically linked to the development of new instruments and tools, and essentially utilitarian inventions such as automatic vocal riding, drum machines, electromechanical keyboards and digital pitch correction have been famously used and abused for creative effect. These advancements have changed the nature of the sound engineering profession from primarily technical to increasingly expressive [12]. Generally, there is economic, technological and artistic merit in exploiting the immense computing power and flexibility that today's digital technology affords, to venture away from the rigid structure of the traditional music production toolset. As more and more industries are starting to consider what implications AI may have (or is already having) for those who work in it, it is opportune to describe the state of the art and offer suggestions about which new tools may become part of the sound engineer's arsenal.

Rapid growth in the quantity of unprocessed audio material has resulted in a similar growth in the engineering tasks and requirements that must be addressed. In audio production for live sound or broadcast one typically has many different sources, each one represented on a separate channel, and they each need to be heard simultaneously. But they could each have been created in different ways, in different environments, with different loudness. Some sources may mask each other, some may be too loud or too quiet, and some may blend in well with the others most of the time, but then have periods where they sound terrible. The final mix should generally have all the sources sound distinct from each other yet contribute to an aesthetically pleasing final product. Achieving this is very labor-intensive, and requires the skills and experience of a professional sound engineer. However, as music content is created by people with a wide variety of backgrounds, there is a need for non-expert audio operators and musicians to be able to achieve a quality mix with minimal effort.

Although production tasks are challenging and technical, much of the initial work follows established rules and best practices. Yet multitrack audio content is still often manipulated 'by hand,' using no computerized signal analysis. This is a time-consuming process, and prone to errors. Only if time and resources permit do sound engineers then refine their choices to produce a mix which best captures an intended style or genre.

1.3.1 Camera Comparison and the 'Instamix'

Professional music production systems offer a wide range of audio effects and processors. But they all require manual manipulation. In effect, as technology has grown, functionality has advanced, but it has not become simpler for the user.

In contrast, the modern digital camera comes with a wide range of intelligent features to assist the user. These include face, scene and motion detection, autofocus and red eye removal. An audio recording or mixing device has none of this. It is essentially deaf. It doesn't listen to the incoming audio, and has no knowledge of the sound scene or of its intended use. It generally lacks the ability to make processing decisions, or to adapt to a different room or a different set of inputs. Instead, the user is forced to either accept poor sound quality, or to do a significant amount of manual processing.

Extending the photography analogy, one could imagine desiring automated parameter adjustments while still controlling the overall style, mimicking a certain technique, or evoking a certain era, which is possible when combining a phone camera with an app like Instagram. Similarly, the amateur music producer or hurried professional could apply a 'filter' corresponding with a certain decade, musical genre, or famous sound engineer, though only if these can be quantified as a function of parameter settings or measurable sound properties. For instance, an inexperienced user could produce a mix that evokes a 'classic rock' sound, a 'Tom Elmhirst' sound, or a '1960s' sound, thus providing a starting point for the novice to achieve a creative goal.

1.3.2 Sound on Sound

Another way to describe the motivation for this work can be found in a 2008 editorial in *Sound on Sound* [13]. The magazine's editor wrote, "There's no reason why a band recording using reasonably conventional instrumentation shouldn't be EQ'd and balanced automatically by advanced DAW software." And that's the point: intelligent software should be able to automate many of the decisions made by a sound engineer.

He also discussed "musicians who'd rather get on with making music than get too deep into engineering," which is an important motivation for us too. One of our goals is to address the needs of the musician who doesn't have the time, expertise or inclination to perform all the audio engineering required. Why should a classical pianist, who has spent the past 20 years mastering their instrument, be required to take a course in engineering, simply to produce a commercial recording?

Finally, there was the comment "[audio interfaces can] come with a 'gain learn' mode... DAWs could optimize their own mixer and plug in gain structure while preserving the same mix balance." That is the specific path that has been taken by many (but not all) IMP techniques: to preserve the choices of the engineer, while automating those other, more standard tasks.

1.3.3 Aims and Objectives

Motivated by these challenges, the research described in this book has aimed to make music production systems more intuitive, more intelligent, higher quality, and capable of autonomous operation. The objectives of much of this research may be described as follows:

- Take care of the technical aspects and physical constraints of music production. For example, limit gain to avoid distortion, or reduce spectral masking of sources to improve intelligibility.
- Simplify complex music production tasks. For example, set some dynamic processing parameters based on features extracted from the source audio.
- Provide intuitive and perceptually relevant interfaces, more directly tailored to the needs of the user.
- Allow for fully autonomous operation, thus enabling amateur musicians and performers to produce content when there are not sufficient resources for a professional sound engineer.

It has also sought to investigate the extent to which these objectives can be achieved. Notably, it explores whether a software system can produce an audio mix of comparable quality to a human generated mix.

1.4 History of the Field

1.4.1 Early Automation in Audio Effects

Many of the classic audio effects were designed as a means to automate, or at least simplify, some aspects of music production.

Early equalizers (EQs) were fixed and integrated into audio receivers and phonograph playback systems, but users and their applications demanded control. Tone controls emerged in the middle of the twentieth century, giving the ability to apply boost or cut to bass and treble [14, 15]. It wasn't until the work of Massenburg [16] that the parametric EQ emerged, allowing control that was more aligned with the needs of audio engineers for smooth yet precise adjustment of the spectrum, attuned to the logarithmic scale (in both frequency and amplitude) of human hearing.

The parametric EQ, though versatile, is not necessarily intuitive. It requires training for use. Reed [9] provided a new interface for EQ that could perhaps be considered the first intelligent EQ assistant. Inductive learning based on user feedback was used to acquire expert skills, enabling semantic terms to be used to control timbral qualities of sound in a context-dependent fashion. Similarly, Mecklenburgh and Loviscach [17] trained a self-organizing map to represent common equalizer settings in a two-dimensional space organized by similarity. However, the space was hand-labeled with descriptions of sounds that the authors considered to be intuitive. A detailed overview of these EQ methods can be found in a recent review paper by Välimäki and Reiss [18].

Automation is more emphasized in the evolution of dynamic range compression. Early use of dynamic range compression was often to fit the signal's amplitude range to that of a recording medium. It also served the purpose of automating the continual riding of faders. But the parameters and use of a compressor are far from intuitive. Tyler [19] provided a simplified version of a compressor, but in the process sacrificed its versatility and much of its adaptive nature. In earlier studies [20, 21], peak and RMS measurements were used to automate the selection of time constants. Automatic make-up gain can be found in some compressor designs, but only as a fixed compensation that does not depend on signal characteristics such as loudness, even though the main purpose of make-up gain is to achieve the same loudness between the uncompressed and compressed signals. More perceptually relevant approaches to dynamic range compression, incorporating some measure of loudness or loudness range, emerged early in the new millennium [22, 23].

However, the biggest leap in intelligent production tools occurred in a multitrack context with the simplest of effects: the fader or level control.

1.4.2 Automatic Microphone Mixing

The idea of automating the audio production process, although relatively unexplored, is not new. It has its origins in automatic microphone mixing. The terminology is used loosely, and automatic microphone mixing is often shortened to automixing. But it refers to a restricted form of automation: automatic microphone level control for speech. In standard speech mixing, it is common practice to open only those microphones that are in use. This reduces noise and maximizes gain before feedback.

Automatic Microphone Mixer Objectives

Dan Dugan [24] stated the basics of microphone mixing and showed that every doubling of the number of microphones reduces the available gain before feedback by 3 dB. Most current designs will restrict the maximum number of microphones to be open regardless of the overall number of microphones in the system. In practice, they generally do not go from a complete 'on' to 'off' state, but instead produce only a 15 dB level change, or some user controllable value [25]. The input channel attenuation setting accomplishes the purpose of optimizing the system gain before feedback, especially when the number of microphones is large.

According to the *Handbook for Sound Engineers* [25], the design objectives of an automatic microphone mixer are:

- Keeps the sound system gain below the threshold of feedback.
- Does not require an operator or sound technician at the controls.
- Does not introduce spurious, undesirable noise or distortion.
- Easy installation, like a conventional mixer.
- Is able to adjust the system status outputs for peripheral equipment control and interface with external control systems.
- Responds only to speech signals and remains mostly unaffected by extraneous background noise signals.
- Activates input channels fast enough that no audible loss of speech occurs.
- Allows more than one talker on the system when required, while still maintaining control over the overall sound gain.

The last three objectives are speech-specific, and require modification for use with musical signals. For music mixing, opening and closing inputs may result in unnatural artifacts.

- Introduction

Existing microphone mixers tend to be part of huge conference systems and often interact with other devices using standard communication protocols. Some of these conference systems have multiple rooms and outputs, which adds complexity to the mixer output stages.

Automatic Microphone Mixer Designs

Early automatic microphone mixer systems were only concerned with automatic level handling and required a significant amount of human interaction during setup to ensure a stable operation [25]. One method for controlling the level is use of an automatic gain control, or AGC. The automatic gain control mixers operate by setting up the quietest active microphone as the reference gain. The microphone has maximum gain before feedback while louder talkers will activate the AGC to reduce the overall level. AGC tends to have similar control parameters to a dynamic range compressor, with attack and release time specified to minimize artifacts. This increases the user complexity of an automatic mixer. In many cases the settings are fixed by the manufacturer, which simplifies their operation but limits their application.

Fixed threshold mixers use a gate to open and close the microphone channel. When the microphone input signal exceeds a given threshold, the gate switches from closed to open. The static threshold has the disadvantage of impeding low-level signals. It may cause artifacts or miss the first sections of a word.

Gain sharing designs [24,26] are based on the premise that the sum of the signal inputs for all active microphones must be below the minimum value that causes feedback. This value is usually set manually by the user.

Directional sensing [27] uses microphones that both capture the source and estimate the ambient noise level. An input is activated only if the source level is significantly above the noise, and only one microphone is opened per active source.

Multivariable dependent mixers take into account both amplitudes and timing of active signals. Initially, all inputs are attenuated. It preserves the relative gains of the speakers since all output gains are the same. If many speakers attempt to talk at the same time the probability of their microphones being open decreases.

Noise-adaptive threshold mixers use dynamic thresholds for each channel. This enables them to distinguish between signals whose frequency content and amplitude is constant, such as air conditioning noise, and signals rapidly changing in frequency and amplitude, such as speech. Other considerations can be added to this design to ensure that a loud talker does not activate multiple inputs.

None of these early systems were designed to work with music. The variable threshold mixer attempted this [24]. It had the ability to adapt the threshold based on a signal from an additional microphone that captures the room noise contributions, a form of adaptive gating. But it was not fully developed, was only capable of controlling levels and was not based on any perceptual attributes.

Ballou [25] summarized the state of the field as follows:

> The operational concepts used in digital automatic microphone mixers have not varied far from the previously described concepts underlying the analog automatic microphone mixers. This is likely to change, but as future digital automatic mixing concepts will be hidden deep within computer code the manufacturers may be unwilling to reveal the details

> of operational breakthroughs; they will likely be kept as close guarded company secrets. New concepts in automatic mixing might only become public if patents are granted or technical papers are presented.

This statement identified a lack of development in this field. It also acknowledged that due to an unwillingness to publish or disclose, many of these advancements are likely to pass unnoticed in the scientific community.

1.4.3 Steps Towards Intelligent Systems

In the early 2000s, research towards IMP systems was in its infancy, partly because digital music production was still an emerging field. Around that time, Pachet et al. [28] presented the idea of maintaining the intentions of the composer and sound engineer while providing the end user with a degree of control. This system provided the user with controllable parameters, constrained to keep aesthetic intention. It was designed to work with pre-recorded material and was unable to deal with live musical sources. It also required human programming.

Katayose et al. [29] proposed a 'mix-down assistant,' which might apply a target profile to source content in order to mimic the approach of a certain engineer. Similarly, Dannenberg et al. [30, 31] described an intelligent audio editor, an auto-adaptive environment where, among other things, tracks would be automatically tuned. It described a system that was capable of maintaining a target, user-defined loudness relationship. It was intended as a tool to automatically time-align and pitch-correct performances, which are auto-adaptive implementations that transcend into the automatic editing realm. But it required a machine-readable score, and hence was highly limited in its application.

Due to the huge growth in audio codecs, research during this period also focused on mixtures of large numbers of audio signals where data compression is achieved by removing sounds that are masked by other sources [32–35]. Although relevant for real-time audio applications where data size is an issue, it is only concerned with mixes produced by summing the sources. And data size is not a key challenge addressed in the music production process.

1.4.4 The Automatic Mixing Revolution

Between 2007 and 2010, Enrique Perez Gonzalez, an experienced sound engineer and music technology researcher, gave new meaning to the term by publishing methods to automatically adjust not just level [39], but also stereo panning of multitrack audio [36], equalization [40], delay and polarity correction [72]. He automated complex mixing processes such as source enhancement within the mix (a multitrack form of mirror equalization) [38] and acoustic feedback prevention [37], and performed the first formal evaluation of an automatic mixing system [43]. To our knowledge, this was the inception of the field as it is known today.

This ushered in a burst of sustained activity in the Automatic Mixing field [73], an important subset of IMP. Figure 1.1 shows a comprehensive but not exclusive overview of published systems or methods to automate mixing subtasks during the ten years beginning with Perez's first automatic mixing paper. Some trends are immediately apparent from this timeline. For instance, machine learning methods seem to be gaining popularity [65–70].

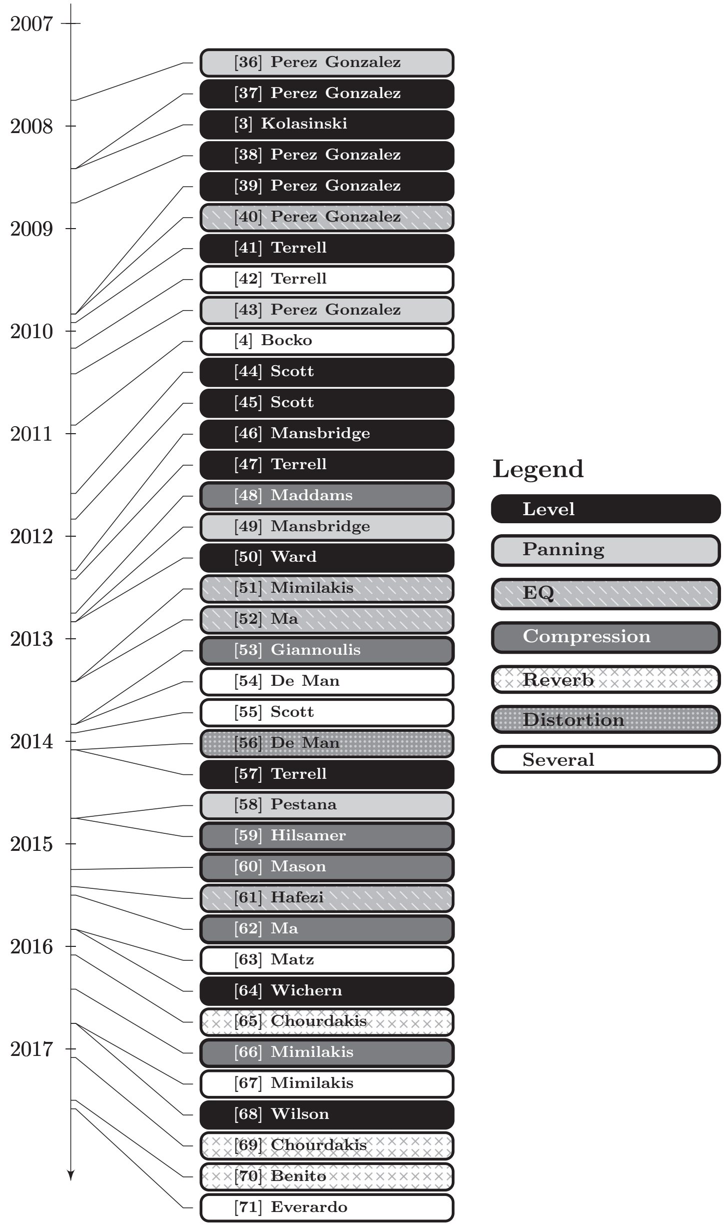


Figure 1.1 Timeline of prior automatic mixing work 2007–2017

Whereas a majority of early automatic mixing systems were concerned with setting levels, recent years have also seen automation of increasingly 'complex' processors such as dynamic range compressors [48, 53, 59, 60, 62, 66] and reverb effects [65, 69, 70]. Research on such systems has additionally inspired several works on furthering understanding of the complex mix process and its perception [12, 74–76].

1.5 Applications and Impact

Currently, most audio processing tools demand manual intervention. For instance, although audio workstations are capable of saving a set of static scenes for later use [77], they lack the ability to take intelligent decisions, such as adapting to different acoustic environments or a different set of inputs.

But with the advent of the technologies described in this book, this is all changing. The purpose of many IMP tools is to reduce the work burden on the engineer, producer or musician, and to explore the extent to which various tasks can be automated. Knowledge about the process of mix engineering has many immediate applications, of which some are explored here.

They range from completely autonomous mixing systems, to more assistive, workflow-enhancing tools. The boundaries between these categories are vague, and most systems can be adapted for less or more user control.

1.5.1 Intelligent Assistants

In 2000, Moorer [78] introduced the concept of an Intelligent Assistant, incorporating psychoacoustic models of loudness and audibility, intended to "take over the mundane aspects of music production, leaving the creative side to the professionals, where it belongs". Adding control over high-level parameters, such as targeted genre, shifts the potential of IMP systems from corrective tools that help obtain a single, allegedly ideal mix, to creative tools offering countless possibilities and the user-friendly parameters to achieve them. This mitigates the risk – and frequent criticism of skeptics – that mixes become generic, as a result of a so-called 'color by numbers' approach lacking creative vision and the intuition of an expert [6].

Sound engineers of varying levels will typically want some degree of control, adjusting a number of parameters of mostly automatic systems. Such a system can quickly provide a starting point for a mix, or reach an acceptable balance during a sound check, like a digital assistant engineer. Even within a single processor, extracting relevant features from the audio and adjusting the chosen preset accordingly would represent a dramatic leap over the static presets commonly found in music production software [3].

Assistants can range from very limited automation – the already common automation of ballistics time constants and make-up gain in dynamic range compressors – to only exposing a handful of controls of a comprehensive mixing system. An example of the latter can be found in online mastering services, where the user often controls as little as a single sonic attribute.

Large audio productions often have hundreds of tracks. Being able to group some of those tracks into an automatic mode, or even automate the grouping, will reduce the effort required by the audio engineer. There is also the possibility of applying this technology to remote mixing applications where latency is too large to be able to interact with all aspects of the mix.

Riders

One attempt to reduce the required effort of the sound engineer while mixing multitrack content is the development of automatic riders. A rider is a type of gain control which continually and smoothly adjusts the gain on a channel to match a given criterion. Recently riders have been devised that, given the desired loudness level of a target channel in relation to the rest of the mix, will compensate for all deviations by raising or lowering the levels of the target [12,79]. The existing riders only work with at most two channels of audio, and thus riders have limited flexibility. All channels other than the target are combined into a single channel, only the target is modified, and only a single loudness curve can be used.

Automating fader riding is a quintessential example of assistive mix technology, as it used to be a tedious and manual job where sometimes several engineers moved a number of faders to change levels dynamically over the course of a song. This task was then automated to some extent by the dynamic range compressor as well as by programmable motorized faders or digital gain automation.

1.5.2 Black Box

In engineering terms, a 'black box' is a system which can only be judged based on its output signals, in relation to the supplied input. In other words, the user does not know what goes on inside, and cannot control it except by modifying the incoming signals. One or more mix tasks could be automated by such a device so that no sound engineer is required to adjust parameters on a live or studio mix [13, 47, 80].

Many of the academic approaches are presented this way, given the appeal of presenting a fully automatic system, although they could be generalized or only partially implemented to give the user more control. The absence of a need – or option – for user interaction is a desired characteristic of complete mixing and mastering solutions, for instance for a small concert venue without sound engineers, a band rehearsal, or a conference PA system.

Live Music Production Tools

Methods have also been proposed for intelligent *live* music production systems. Some of these tools take advantage of mixing board recall functionality and add autonomous decisions to the mixing process. Much of the research described here was for live sound applications. But to our knowledge, no commercially available mixing device is yet capable of equalizing and organizing the gain structure while taking care of acoustic or technical constraints.

1.5.3 Interfaces

Another class of IMP tools, complementary to Automatic Mixing in the strict sense, comprises more or less traditional processors controlled in novel ways. For instance, a regular equalizer can be controlled with more semantic and perceptually motivated parameters, such as 'warm', 'crisp' and 'full' [81, 375], increasing accessibility towards novices and enhancing the creative flow of professionals.

Deviating from the usual division of labor among signal processing units, the control of a single high-level percept can be achieved by a combination of equalization, dynamic range compression, harmonic distortion or spatial processing. Such intuitive interfaces are likely to speed up music production tasks compared to traditional tools, but also facilitate new ways of working and spur creativity. In the pro-audio realm, adaptive signal processing and novel interfaces allow users to try out complex settings quickly [17].

Already, music software manufacturers are releasing products where the user controls complex processing by adjusting as little as one parameter. New research is needed to validate these relationships, uncover others, and confirm to what extent they hold across different regions and genres.

1.5.4 Metering and Diagnostics

Intelligent metering constitutes another possible class of systems built on this new information, taking the omnipresent loudness meters, spectral analyzers, and goniometers a step further, towards more semantic, mix-level alerts [410]. For instance, the operator can be warned when the overall reverb level is high [82] or the spectral contour too 'boxy.' By defining these high-level attributes as a function of measurable quantities, mix diagnostics become more useful and accessible to both experts and laymen. Such applications also present opportunities for education, where aspiring mix engineers can be informed of which parameter settings are considered extreme. Once these perceptually informed issues have been identified, a feedback loop could adjust parameters until the problem is mitigated, for instance turning the reverberator level up or down until high-level attribute 'reverb amount' enters a predefined range.

1.5.5 Interactive Audio

Growth in demand for video games and their ever-increasing audio processing requirements makes Intelligent Audio Production for games a promising research area. In game or interactive audio, hundreds of ever changing audio stems need to be prioritized and mixed in real time to enhance the user experience. Virtual reality and game audio often have dynamic environments where users may change their location with respect to the sources, and sources are then rendered to give the appropriate spatial characteristics [28]. This suggests further constraints; some sources have a predetermined spatial position and others can be adjusted. Any automation or streamlining of this process could be beneficial, especially since the automation is not 'competing' with a professional, creative sound engineer. That is, the challenge is to either perform Intelligent Audio Production, or not perform anything except a simple sum of sources with some basic rules applied.

Audio Augmented Reality for Interactive Soundwalks, Sound Art and Music Delivery

Dafna Naphtali and Richard Rodkin

14.1. Introduction

Composers have long sought to utilize public spaces as a means of enhancing audience experience, either by providing a visceral and unique context for performance or by using the environment itself as a collaborator enlisted to influence a piece's outcome. Historically, the desire for this type of engagement with interesting acoustic spaces goes back centuries, if not millennia, to everything from mountain-top yodeling, to sixteenth-century Venetian polychoral music, early church music and antiphonal chants, and in many cases influenced the architecture itself. As such, many historical and contemporary examples of "site-specific" compositions were written for specific acoustic spaces: either natural environments or man-made spaces in interaction with architectural forms or urban design. In some cases, works include the movement of performers or listeners within those spaces, sometimes even turning the spaces themselves into virtual musical instruments.

With advances in technology, the methods for unifying musical experience with place have also become more varied and sophisticated. Now listeners can not only engage with the acoustics of a space, but they can also interact with the location itself, drawing from its inherent symbolism and ambience, thereby heightening the emotional response even further. Now called "location-based music" "location-based sound," or "locative media," the approach embodies many different styles of music and activities, yet it always requires the use of a specific physical space, a method for tracking the listener within that space, and a set of rules for how and when the music is delivered based on the listener's position and/or interactions.

As a subset of the locative media category, audio augmented reality (AAR) is not necessarily new, but the advent of the smartphone and its peripheral technologies have presented an opportunity to express and share these personal and social experiences—even one's life story—in the most engaging and imaginative ways possible, and with far wider audiences. Having collaborated on an AAR installation in New York's Washington Square Park, the authors will reference both the piece—"Walkie Talkie Dream Angles," by Dafna Naphtali—and the system with which it was constructed—Richard Rodkin's U-GRUVE AR—throughout this chapter. The discussion will first examine the available technical components needed to create an AAR system, as well as the decisions and considerations necessary for choosing the appropriate components given certain sets of requirements. The focus will then shift to explore the various compositional techniques and approaches that must be thought through to ensure the most engaging and rewarding experience for the participant.

As a compositional practice, the body of work comprising the locative media category is quite expansive. These projects and experiences—ranging from installations to performances to soundwalks—often operate on multiple levels, as something in between self-guided performances, listening experiences and games. While some are cited throughout the chapter, it is impossible to include them all; a section at the end of the chapter has therefore been provided, which includes a more complete listing of AAR projects and related readings.

14.2. Audio Augmented Reality Defined

It is important to establish a clear definition of "audio augmented reality" early in this discussion, since, as mentioned above, the history and technique for delivering location-based sound art is long and varied. To do this the underlying term "augmented reality" must first be defined.

Augmented reality (AR) has largely been thought of as a visual medium, in which a mobile device—either hand-held or head-mounted—presents an additional, virtual layer of contextual information over top of the user's real-world experience, consisting of either two-dimensional text-and-image display, or three-dimensional imagery that appears to have been inserted directly into his or her surroundings.

AAR employs the same principle as AR, but the overlaid information is presented in audio form. In certain frameworks, AAR strictly refers to non-musical, non-artistic audio information, for example in systems designed to aid visually challenged people, where audio cues can provide wayfinding, traffic and basic point-of-interest information. There is also a branch of AAR focused chiefly on self-guided touring, in which historical information or stories are presented by a narrator as the participant walks through a given area. And while spoken word certainly has a place within artistic performance, narrated tours are not the focus of our discussion.

For this discussion, AAR refers to the real-time delivery of audio artifacts, whether prerecorded, generated or live, in response to the participant's environment, and for the purpose of creating an artistic, performative experience for and with the participant.

Finally, to further clarify, AR/AAR methodology falls into two general types: location-driven and environment-driven. Location-driven systems require a method of detecting the participant's presence, either through a tracking system or through marker/object detection, and then triggering audio events per the composer's rules. An environment-driven system can ostensibly be used anywhere, regardless of specific location, and instead responds to direct stimuli from the participant's surrounding environment (weather, time of day, traffic conditions, environmental sounds, etc.). The discussion that follows will be organized around these two general approaches to AR/AAR.

14.3. Technological Approaches to AAR

14.3.1. Elements of an AAR System

Though not prescribing a specific toolset or software platform upon which to build an AAR system, this discussion focuses on the components and associated factors that need to be considered in the planning and preparation of any viable AAR project. As a basis of comparison, we'll be referencing U-GRUVE AR, as mentioned above, which has been successfully used by the authors for multiple projects over the past several years.

14.3.1.1. System Components

It is useful to review the components comprising any AAR system, as this understanding will enable system designers to clearly traverse the decision tree needed for defining functional requirements. While implementations will vary, all AAR systems consist of: (1) at least one sensor for either tracking the participant or for detecting a specific environmental input, (2) a method for controlling sensor inputs, (3) a method for mapping control events to the composer's rules and conditions and (4) the physical device used by the participant, which houses the sensors, software and playback mechanism.

14.3.1.2. Sensors

AR is a predominantly mobile technology, as its real value lies in its ability to "read" any context and provide the participant with some type of relevant, real-time feedback. The most critical components, therefore, are the sensors used for establishing this context. Sensors can range widely in terms of what they can detect and how, but all are generally used for either tracking the participant's position or for capturing and manipulating environmental elements within the participant's proximity. Tracking systems consist of two subcomponents—the tracking device worn by the person being tracked, and the central tracker used to monitor the tracking device(s).

14.3.1.3. Control Methods

While sensors are used to connect the AAR system with the participant, nothing will happen without a method for invoking the audio elements, either through a system of triggers or by using a data stream to apply continuous control.

Trigger-based controls fire events when specified thresholds are crossed. Location-driven systems using GPS as the primary sensor naturally lend themselves to zone-based trigger methods. Once the zones have been laid out, it is then just a matter of attaching the desired audio artifacts and mapping them to the ruleset defined by the composer.

Camera-based systems use fiducial markers as the trigger method. With this approach, a "match" on an image or pattern can be used to fire an event, in exactly the same way that crossing into a zone would.

For environment-driven systems, triggers are more data-oriented, similar to the GPS example above, but obviously unrelated to the user's location. For example, a thermal sensor could be used to read the temperature around the participant, with discrete temperature values used as trigger points. With a high enough resolution, the sensor could be set up to record ranges of temperature values, which could allow for continuous control over the audio.

As an alternative to triggers, incoming data can be used to apply continuous control to various audio parameters. In a location-driven system, for example, a stream of GPS values could be used to manipulate the resonance on a filter. Similarly, an environment-driven system might also use data such as speed, heading or orientation to affect the sound in some way.

14.3.1.4. Rules and Conditions

Whatever the control method, the heart of the interaction lies with the rules and conditions that are invoked once a control event is registered within the system. This could be as simple as "Play Sound X while the participant stays within Zone A" or something as sophisticated as reading weather data and assigning a randomly selected audio clip. In *Walkie*

Talkie Dream Angles (WTDA) certain sets of sounds were programmed to play depending on time of day, with no sound in a set being selected more than once. These rules are executed via code, and therefore really can run the gamut of complexity, limited only by the composer's imagination and programming skills. That said, a programming or scripting language must be selected to establish and manage the processing between events, rules and audio responses.

14.3.1.5. The Delivery Mechanism

The device that the participant will use to engage with the system is key to the overall success of the application. It must not only securely house the selected sensor(s), but also must be able to run the software that logs and maps the control events against the rules and then both access and play the requested audio. This could be something as convenient as a smartphone, or something completely custom built, such as a vest embedded with wearable technology.

Having now defined the core set of components needed to construct an AAR system, the project goals and requirements can now be assessed against the range of available solutions.

14.3.2. Selecting the Right Technology

Before evaluating any AAR technologies, some basic functional requirements must be thought through, as they will determine the participant's core experience. These requirements, of course, should be primarily driven by the artistic intent, though sometimes compromises may be necessary due to cost, feasibility or skill limitations.

First, which mode of AAR will be used—location-driven or environment-driven? Or both? Next, is the chosen location indoors or outdoors? Then, what kind of device will participants use—one that they already own and bring with them, like a smartphone, or one that will be custom built and loaned out? Will the experience be moderated by an attendant, or will participants be able to operate the piece autonomously? Will there be only one composition available, or will there be a variety? There are myriad other aspects to examine, but let's discuss these core questions as they will all influence the construction of the system.

14.3.2.1. Location-Driven or Environment-Driven? (Or Both?)

As mentioned earlier, AAR experiences utilize either a tracking method for determining the participant's physical location or a set of sensors for collecting information from the participant's immediate surroundings, which are then processed and matched against the composer's ruleset. This is not to say that the two methods are mutually exclusive. It could be quite feasible to implement both a tracking system and an array of sensors for scanning the participant's environment. But the decision needs to be made very early on in the planning process, as it will drive subsequent decisions on both locale and form factor for the delivery device.

14.3.2.2. The Tracking System—Outdoors Versus Indoors

For location-driven AAR experiences, first decide whether the system will be used indoors or outdoors, since the tracking technologies available are largely governed by this factor. In some cases functional limitations will determine the system (GPS, for example, does not work well at all indoors), while in others the project directive itself may drive the decision—if the project is for a museum or another type of indoor installation, then the choice has already been made, and the options are predetermined.

14.3.2.2.1. Outdoor Systems

GPS is the most common form of tracking for outdoor systems. As a near-ubiquitous technology found in most mobile devices and motor vehicles, the costs are relatively low, and there is an abundance of documentation and developer APIs available. There are also very low-cost, micro-module units available, which could be embedded in wearable devices in any number of ways.

Despite its convenience and general accuracy when used for wayfinding, consumer-grade GPS does come with a few caveats. First, positioning is only accurate to ten meters, and while it is wholly possible to achieve better performance in actual practice, artists should plan their pieces according to the ten-meter rule. Secondly, GPS signals can drift from moment to moment, which could result in unpredictable control events. Since these fluctuations aren't handled automatically, this needs to be custom coded by the developer. Finally, GPS signal deteriorates markedly when line-of-sight is broken between the receiver and the satellite. Conditions in urban environments where dense groupings of tall buildings can potentially block satellite signals will therefore need to be further mitigated. Again, software fail-safes will need to be implemented ("If no signal, play sound X," for example), so relevant planning cycles should be allotted for this.

Another option for outdoor tracking is the standard camera-based AR approach for detecting fiducial markers. As this is the era of ARKit and ARCore, the most obvious approach would be to tap into the potential of those systems. The inherent drawback to using a marker-based system, however, is that the markers need to be where they are expected to be at all times. It is highly unlikely that markers placed in a public park will be there the next day. That is not to say there aren't workarounds—projects sponsored by a group that manages the outdoor space may find it possible to construct and position something more durable.

Also consider that the participant must point the camera at the marker for it to be recognized. This means that (1) the participant must always be on the lookout for markers, which may distract from the overall experience, and (2) the participant is forced to always have their device in hand, which may get tiresome or, again, distracting.

An alternative camera-based approach utilizes a form of computer vision in which the camera's inputs are continually scanned and compared to a library of images, and when a match is found, the trigger event is fired. Even more sophisticated computer-vision algorithms can be applied, as with the recently debuted *Play the City*[1] project (nocomputer, 2018). In this case, an automobile was outfitted with a sophisticated rig consisting of cameras and the requisite processor to handle the real-time analysis of potential "targets" while the car was in motion. This is a highly specialized example, but in general, any solution of this nature will require more computational muscle, so it may be prohibitive in cost, programming sophistication and form factor.

Similar to fiducial markers, though a distinctly different technology, are iBeacons. These are small devices that use BlueTooth to detect the presence of a specific app on a smartphone, and deliver programmed content based on the user's proximity to the beacon. iBeacons are mostly used in retail and museum settings, but in certain controlled situations, could be used outdoors, where they would have to be securely mounted out of sight, in a window or kiosk of some sort. While not as flexible as GPS, it is wholly possible to utilize the beacon's three levels of proximity detection (close, mid-range and far) as a means of triggering audio events managed by a mobile app.

In the development of U-GRUVE AR, it was clear from the outset that it would be an outdoor system, as the primary goal was to provide interactive musical backdrops for public spaces such as plazas, parks, gardens and rooftops. Additionally, the goal was to be able to reach as large of an audience as possible, so creating a mobile app for smartphones also made the most sense. The mobile app approach also satisfied our requirement for a self-guided and untimed experience. Given that a GPS sensor is standard equipment in all smartphones, the decision to use GPS for tracking was more or less a foregone conclusion.

14.3.2.2.2. Indoor Systems

For many artists, indoor installations are very appealing because the environment can be controlled and protected, allowing for artistic possibilities that wouldn't otherwise be possible in open, public spaces. While solutions such as iBeacons and fiducial markers will work just as well indoors as they will outdoors, they also bring the same drawbacks, namely, that the participant must actively use their smartphone during the experience.

Alternatively, there are a plethora of indoor positioning technologies that can be used to track individual participants. The most common systems are radio (RFID, BlueTooth), optical (infrared, visible light) and magnetic, among others.[2] It would be far beyond the scope of this discussion to provide an in-depth analysis of each of these technologies, but the key takeaway is that to implement of any of these systems requires custom design, installation and testing. If an installation is touring from site to site, it will need to be fully installed and tested at each new location.

Additionally, most of these technologies require that a physical tracker be either attached to or carried by the participant in some way. In the case of RFID and InfraRed (IR), for example, the user wears a small badge whose signal is read by a central receiver, so this aspect will directly impact the form factor of the delivery device.

14.3.2.3. The Delivery Device

Ultimately, the decisions made about the type of AAR to be used—location-driven or environment-driven—and whether the work will be indoors or outdoors will inherently influence the delivery device that participants will use.

14.3.2.3.1. Handheld

Of course there are other basic factors to consider, in terms of artistic vision and general goals for the work. For example, U-GRUVE AR was designed from the outset to construct a veritable platform that would make the works of many artists, at many locations, available to as many participants as possible. Additionally, the participant would be able to experience a piece without mediation or time limit. Given those requirements, the smartphone, paired with a good set of binaural headphones, was the obvious choice.

There are other options, however. In museum and gallery settings, it's quite common for participants to check-out tablets for self-guided tours, so it wouldn't be unusual to opt for this approach in an indoor setting. Depending on the system, it may also be necessary to design and build a custom device using Arduino or Raspberry Pi.

14.3.2.3.2. Ridable

While the general assumption is that participants will be walking while engaging with an AAR experience, there is precedence for what might be called a "sound drive," for lack of a better term. In one early example of location-driven AAR, composers Jesse Stiles and Melissa St. Pierre outfitted a car with a GPS receiver and a computer to create the "GPS Beatmap,"[3] which was demonstrated on the Bonneville Salt Flats in 2006. To experience the piece the participant would have to drive the car, or be driven, through various zones to trigger the audio, which was transmitted from loudspeakers mounted on the car's roof, as well as through the car's internal speaker system.

The *Play the City* experience mentioned earlier also used a car as the device, with the cameras mounted outside the car, and an AR "viewer" overlaid on top of one of the rear passenger windows. In this case, the participant is always a passenger in the moving vehicle, and can therefore focus solely on the musical experience.

When testing U-GRUVE AR's first version, sound-rides for cyclists were considered a core offering, but in light of concerns over potential accidents and legal issues, this approach was shelved. However, sound artist Kaffe Matthews has found a way to circumvent these issues via her *Sonic Bike* project,[4] which features a self-produced, highly customized bike that is outfitted with speakers, through which GPS-triggered musical events are played—similar to how U-GRUVE works, but played exclusively through the speakers, thereby avoiding the risks incurred by masking environmental sounds with headphones. In addition to using location as a trigger for controlling playback, the output is also influenced by the biker's pedaling speed. The project, which began in 2008, continues to be extended and advanced at the Bicrophonic Research Institute, led by Matthews and Dave Griffiths, and now includes the production of the *Sonic Kayak* and the *Sensory Bike*.[5]

14.3.2.3.3. Wearable

Before the advent of the smartphone, AR applications usually required the construction of some sort of customized wearable setup, utilizing either a vest laden with wiring and sensors and/or a helmet containing a camera or GPS receiver. While technologies have advanced and components have gotten smaller and lighter, wearable systems could still be an attractive solution, especially with the recent trends in the fashion tech industry. Here again, as with custom-built handheld devices, production, distribution and maintenance all need to be considered.

14.3.2.4. The Software Platform

Closely linked to the decision on the delivery device, the choice of software platform for pulling everything together—that it is, processing sensor data, setting control methods, defining rulesets and assigning the correlating sound events—is critical, as it will impact factors like cost, extensibility and maintenance. Again, this choice is guided by the overall scope of the project, in that production of a one-off experience may not require the coding rigor or depth of features that is often necessary in a more widely distributed product.

Just as there are integrated development environments (IDEs) such as XCode and Android Studio, there are also complete platforms to consider, such as the game engines used for developing today's vast array of mobile apps. Often these packages, such as the Unity Game Engine and Unreal Engine, come with features for integrating smartphone capabilities, audio processing and scripting logic, as well as "one-click" deployment to various gaming platforms. That feature alone can't be overstated enough, as it potentially spares the developer from having to recode versions of the application for each deployment target.

14.3.3. Audio Generation/Playback

14.3.3.1. Prerecorded Audio Clips

The most straightforward approach to playing audio is to draw from a set of prerecorded audio clips. Using clips ensures an overall audio consistency, barring any parameterized variations in volume or effects handling. While not offering the flexibility of pure synthesis, there are many techniques to employ for varying the sound, such as programmatic playback of clip segments, randomization of clip sequences, and dynamic looping, to name a few.

Regardless of platform, because AR is delivered through mobile media, clip size and resolution should always be kept in mind during the production process. Whatever device format is chosen, a strategy and method for managing the loading in and out of audio files is needed. In some cases it may make sense to preload all of the audio clips, while in others, a more sophisticated plan may be needed, in which clips are loaded dynamically based on the participants specific location or proximity to certain zones. These strategies should also take into account factors such as the participant's connectivity, be it cellular or wi-fi, and average storage capacity on the expected delivery device.

14.3.3.2. Sound Synthesis

Integrating an actual synthesizer into the sonic palette introduces another level of variation and expression, as participants now have control over not just the starting and stopping of a clip, but over the very elements of the sound itself, from the shape of the filter envelope to an oscillator's waveform.

There are a number of options for harnessing software synthesis in mobile contexts. For the iOS platform, AudioKit offers a comprehensive API for harnessing Apple's AudioUnit SDK, allowing for relatively straightforward implementation of synthesizers and audio processors. Within a narrower scope, another option, which is also Android-compatible, is LibPD, which takes the well-known PureData synth engine and embeds within the platform's framework. This package was developed by Peter Brinkmann, and has been documented extensively through his book *Making Musical Apps*.[6]

If the device is custom built, there are many options for integrating both hardware and software synthesizers with microcomputers such as the Arduino or Raspberry Pi. Synths like Fluxamasynth,[7] which consist of a hardware "shield" component that is controlled via a pre-existing software library, offer a relatively easy path to this realm of music making.

Despite these options, the expressive flexibility that comes with native synth engines also carries one significant caveat, which goes more to understanding the audience. While it may seem like a great idea to give participants complete control over the sound, if they aren't necessarily experienced with sound synthesis, care must be taken to make the experience foolproof so that unpredictable outcomes, such as volume loss or resonance spikes, never happen. As soon as the participant feels like she or he has lost control of instrument, she or he will lose interest in the performance as well. Those with more music-making experience may also have more patience for experimenting with new instruments, so it is imperative that the target audience is clearly understood before committing to this type of sound generation.

14.3.3.3. Live Sampling/Live Sound Processing

While certainly not limited to environmental-based AAR, sampling sounds around the participant and processing them in real time is a hallmark technique used in non-location-specific experiences. In fact, this was the signature feature of one of the first AAR apps on the market, RJDJ, and its sister app, Dimensions, released for iOS in 2008. Seeking to establish a genre that co-founder Michael Breidenbruecker referred to as "reactive music," RJDJ's premise was to "sense" the sounds around the participant via the smartphone's microphone, and then route the audio through any number of filters and effects for live processing, which were collectively called a "scene." Whether the participant would actively speak into the microphone or a bus would go whizzing by, the audio would be processed per the rules of a given scene, which could be either preset or defined by the user and controlled further using other built-in sensors (camera, accelerometer, GPS).

It should also be noted that RJDJ didn't just passively receive and process audio. Participants could also manipulate sounds by moving the iPhone itself, waving it or dragging it through the air like a wand. In this case, the phone's built-in accelerometer and gyroscope were used as continuous controllers that applied various effects on the incoming audio signal.

While RJDJ has left the scene, there are audio processing alternatives for mobile platforms, such as SuperPowered,[8] which offers a very robust solution for both iOS and Android. In addition to its low-latency processing muscle, SuperPowered also features real-time pitch shifting and time stretching, a capability that was once a luxury on the desktop and is now available in mobile devices, significantly expanding compositional and performance possibilities.

14.3.4. General Considerations

14.3.4.1. Synchronized Versus Non-synchronized Playback

Some musical styles may warrant that tracks play in sync. We're mentioning it here since, depending on which software platform is chosen, this capability is not necessarily a given. Chances are, the developer will need to either write or purchase some code to enable this feature.

14.3.4.2. 3D Spatialization

"3D spatialization" is the term used for giving a sound attributes so that it is perceived as occupying a distinct location within a three-dimensional environment. This is achieved by attaching the sound to an object that occupies a specific location, and then applying various audio parameters that represent that location in relation to the participant's position. Effectively, this means that as the participant gets closer to the object, the associated sound will get louder, and if he or she approaches the object from the left, the sound will be panned toward the right side of the stereo field.

While normally used in video game settings to help the player accurately orient themselves within the game, 3D spatialization can also be effectively applied in much the same manner as well to help focus the participant's attention on specific objects. For example, an audio clip could be "attached" to a statue in a park to emphasize its significance and otherwise bring it to life. Alternatively, one could create the sounds of a virtual street musician and locate the performance in a specific place to create the illusion of presence.

14.3.5. User Interface

The final component to consider in designing an AAR system is the user interface. This is the "final" component because its format will largely be driven by previous decisions around the delivery device, its form factor and the manner in which it is going to be used. The important thing is that this aspect not be overlooked or underestimated, as it can be crucial to the project's success. If users become frustrated with merely accessing and managing the piece's content, they simply will turn it off and not use it.

There are generally three types of user interfaces to consider, none of which are necessarily mutually exclusive: the graphical user interface (GUI), the hardware interface and the audio user interface (AUI).

14.3.5.1. Graphical User Interfaces

In the realm of mobile media, the GUI is the most prevalent user interface, though obviously not without the support of a minimal set of hardware buttons [for unlocking the phone, controlling volume and powering on/off]. If the delivery device is a smartphone, then at least one screen will need to be designed for accessing the app. From there, the depth and complexity is entirely dependent on what the app is meant to do.

In the case of U-GRUVE AR, which actually consists of multiple locations, each of which may have multiple pieces available by multiple artists, the GUI is a critical component not just for experiencing a piece, but for navigating to and selecting it. The conventional navigation patterns most commonly associated with music listening apps were used, such as drill-down lists leading to landing screens for each location, with each of those consisting of the location's description and track list. The side-menu pattern was also utilized to present information about the app, the locations and the composers.

The design also included a screen once a piece was launched that would display a graphical representation of the participant over top of the "MixMap" for that piece, similar to the blue dot that appears when using a mobile map application. Within U-GRUVE AR the MixMap is the visual depiction of the various zones to which audio artifacts are assigned (see Figures 14.1 and 14.2). While the MixMap serves several utilitarian purposes during the production process, it is also useful to participants for establishing their orientation within the environment and for helping them see where they might find various sounds.

14.3.5.2. Hardware Interfaces

Custom devices may utilize a GUI or AUI, though by virtue of being custom built, the standard GUI may be bypassed altogether in lieu of a purely

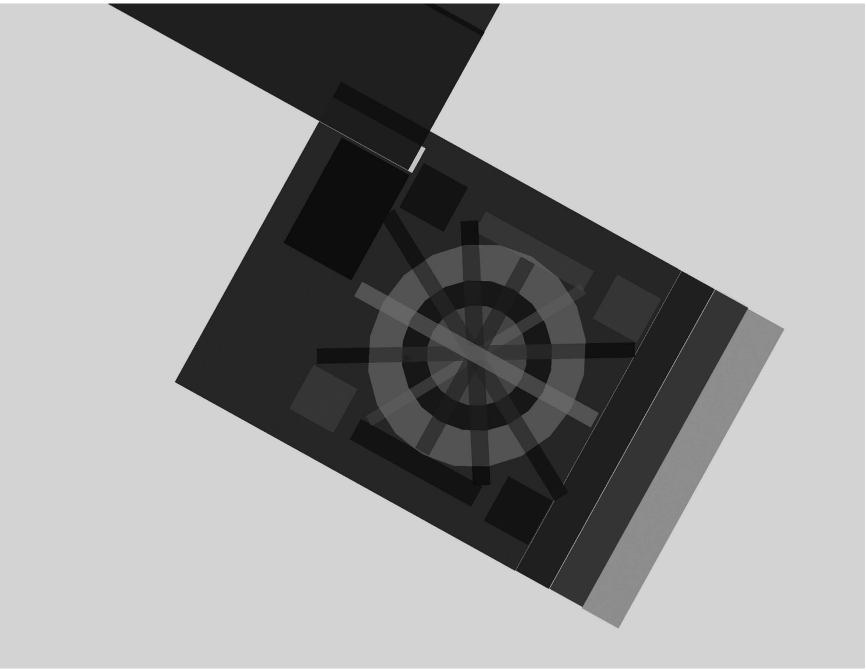


Figure 14.1 MixMap for *Orela* (2016), by Barbara J. Weber, Lincoln Center Plaza, New York.

hardware interface. This may consist of physical knobs, sliders, toggles—anything, depending on project requirements—and in some cases this may offer an advantage, depending on who the users are. If, for example, the experience is for preschool children, they may engage more readily with a large button than with a flat screen. The key, though, as with any UI, is to maintain a level of simplicity and avoid "cockpit syndrome" in which users become overwhelmed with too many controls.

14.3.5.3. Audio User Interface

Oddly enough, with all of the recent work in what might be called "clinical" AAR—that being the provision of an audio interface for informational and wayfinding purposes (sometimes referred to as "sonification"), the term AUI has not really found its way into the formal lexicon. While a GUI is normally used to help navigate the application's screen-level content, an AUI would be used as an assistive layer for helping users understand boundaries, locations of new sounds, or other in-experience information that's been deemed relevant for the participant to know. Creating a useful

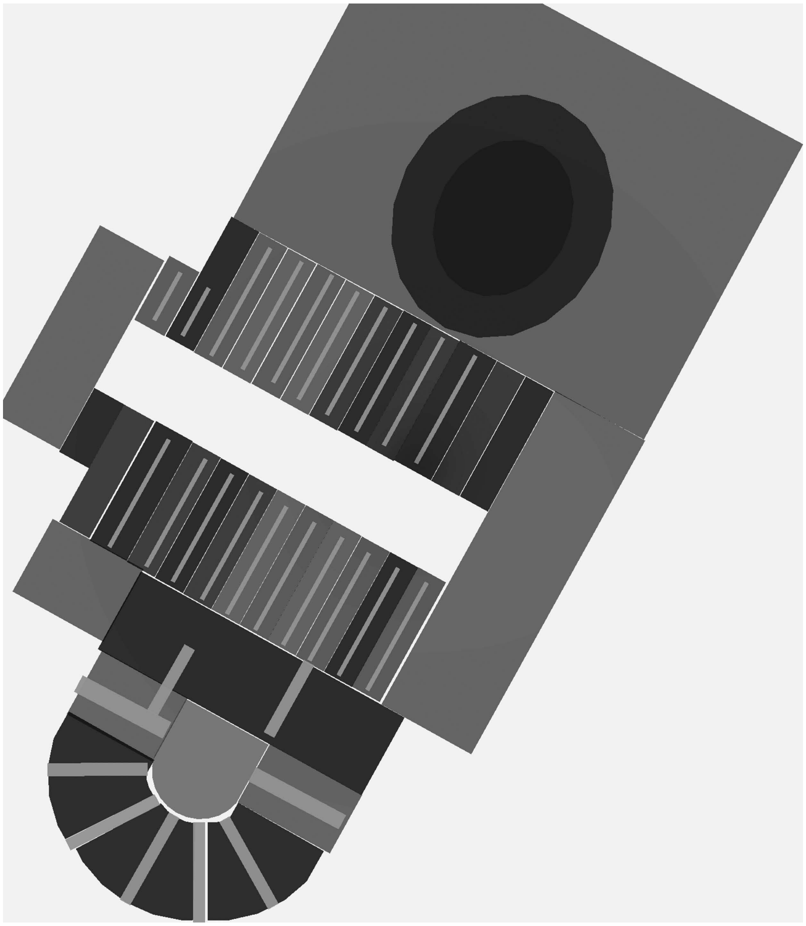


Figure 14.2 MixMap for *Florambula* (2017), by Barbara J. Weber, Conservatory Garden, Central Park, New York.

set of "earcons" could prove very valuable but would also require some degree of usability testing, as interpretations of what a sound means can vary widely, just as with their visual counterparts, icons. Care should also be taken to choose tones that are at once not alarming nor disruptive, yet are still markedly distinct from the actual audio content of the piece.

All of that said, another perfectly valid choice for a UI would be to not create any interface at all, and rather just provide a "start and go" experience in which the participant simply puts on headphones, touches a single button to start the piece, and tucks the device away while listening without distraction of any kind.

14.4. Compositional Considerations

The creative considerations for a composer working with AAR are many and profound. The possible issues, solutions, modes of expressions and content are as multifaceted and vast as the world of music and sound art itself. Successful projects will use interactive or adaptive methods, many drawn from game music composing, sound installations, data sonification and other forms of interactive media. The innumerable ideas about designing interactivity should be considered as independent of the compositional styles used for writing the music.

Using AAR with GPS location affords an opportunity to add several layers of sound and meaning to a composition. Among the possibilities are musical underscoring, sounds giving historical/social context, audio cues and sound elements that imbue the listener with agency to interact directly with and possibly even "play" the music themselves through their movements. Below are outlined ideas and strategies drawn from the authors' own work, interspersed with a discussion of projects by others.

WTDA is a location-driven, electroacoustic soundwalk composed by Naphtali in collaboration with Rodkin using his U-GRUVE AR audio augmented reality platform. Created for NoiseGate, an acoustic ecology festival in New York City in 2016, WTDA celebrates the soundscape of Greenwich Village's historically and sonically rich Washington Square Park, a ten-acre public park in the heart of Greenwich Village.[9] Washington Square Park has been a personal, cultural and political touchstone for many generations of New Yorkers, and a must-see destination for the many tourists who visit each day.

WTDA uses environmental sounds, mostly recorded on-site, and electroacoustically processed compositions created from those sounds to evoke augmented, imagined and/or historical contexts, layering them with percussion and vocal work by the composer. The piece was designed not only to share the personal expression of the composer's experiences in the park (as a lifelong New Yorker) but also to cause the listener to explore and encounter sonically interesting corners of the park in a new way, to underscore the loss of quiet, past and future sounds—a many-pronged approach taken to creating the work, and strategies about site-specific listening, interactivity and attention span.

Although the piece is driven by location, not the environment, it relies on the sonic environment of the park for most of the sound sources. It is assumed that the collages and electroacoustic pieces will be heard, overlaid and aleatorically interwoven with actual live sounds a listener would hear and overhear as they walk through the park. The work is heavily inspired by the work of John Cage, Pauline Oliveros, Hildegard Westerkamp and others who have changed our ideas about the nature of listening, and on a more philosophical/aesthetic level, by the concept of psychogeography[10] and "socio-emotional" concepts about space/sound/emotion and memory.

14.4.1. Planning/Mapping

In the planning for an AAR piece, the first concern will usually be the mapping of the geographical area for the piece into regions or zones. In the case of U-GRUVE, it was easiest to work with an interactive map such as Google Maps[11] to create GPS boundaries ("zones"). Since U-GRUVE AR uses this sort of map directly (called a "MixMap"), drawing out the zones was not only useful for visualizing the sonic structure, but also ensured that the planned arrangement would be translated accurately during the programming phase of the project. The zones may be very large or small, and may be underscored or merely used as triggers to help determine information about how the listener is traversing the area (direction, speed, activity).

The composer will next need to decide what the character is of each zone compositionally, and how the information about a participant's movement is to be used. This is obviously not an ironclad rule, as some composers may do all of the above at the same time on location and others may write the music and then decide how to trigger it later. The relationship between the interactive elements and the audio content itself will likely be different for each composer and each location, as is seen in work by Jesse Stiles, Kaffe Matthews and many other projects listed at the end of this chapter.

For WTDA, Washington Square Park was divided up into multiple overlapping zones/layers, in an approach determined by several factors. Inspired by acoustic ecology "soundwalks" in the tradition of Hildegard Westerkamp, R. Murray Schafer and others, the composer's initial task was to consider the environmental/sonic identity (and actual sounds) in various areas of the park. Each zone related to the physical layout and/or visual attributes of that location, and to the sonic identity that distinguished it from other areas of the park. Other zones were created because they represented inaccessible areas in or near the park whose sounds could be brought out into the public sphere. Still other zones related to emotional/physical memories: of sounds that were either once in the park (in the composer's lifetime or in a more historical context), or sounds that were whimsically imagined could, or should, be there. A unifying factor was that all the zones were viscerally demarcated through some combination of these parameters. It was not simply musical underscoring of the park, but rather a denoting of sounds in a way deeply related to psychogeography and dérive, an attitude toward urban exploration first defined by Guy Debord and Parisian Situationists in 1955. In all of these zones, Naphtali created collages and electronic/

Figure 14.3 MixMap for *Walkie Talkie Dream Angles* (2016), by Dafna Naphtali, Washington Square Park, New York.

electroacoustic pieces based on local recordings and sound effects. Both the acoustic collages and the alternative audio files are triggered from a randomized list. The composer has a long interest in the intersection between birds and technology, and in WTDA one can hear both (in parts of the park where actual construction noise is loud by day, one can still hear birds. At night the virtual construction noises continue to be heard.)

1. *Accounting for participant's perception of time*. Using time as a control factor can be an overlooked way to increase variety and contextualize a piece, and keep it interesting for longer.

   a. *Keep moving*. The tempo and underlying rhythms can be useful for sending cues as to whether to walk or linger in a particular zone. In WTDA, voice and percussion loops, fragments of a wordless

Zone	Event Trigger		Audio Clip Name	Stereo/Mono	Num Loops	init Vol (0-100)	Is 3D (Y/N)	3D Sound-Volume		Delay		
	On	Off						Edge Vol (0-100)	Center Vol (0-100)	Time (ms)	Feedback (0-1.0)	Dry (
WSQ	could loop 1-2 times		VocalDrone2.aif	Stereo	1 or 2 then next in list or randomized							
WSQ	could loop 1-2 times		VocalDrone2a.aif	Stereo	1 or 2 then next							
WSQ	could loop 1-2 times		VocalDrone2b.aif	Stereo	1 or 2 then next							
WSQ	could loop 1-2 times		VocalDrone2c.aif	Stereo	1 or 2 then next							
WSQ	could loop 1-2 times		VocalDrone2d.aif	Stereo	1 or 2 then next							
WSQ	could loop 1-2 times		VocalDrone3a.aif	Stereo	1 or 2 then next							
WSQ			VocalDrone3b.aif	Stereo	longer use only at full park and when in no other zone or on way to Harvestworks							
WSQ			VocalDrone3c.aif	Stereo	longer use only at full park and when in no other zone or on way to Harvestworks							
WSQ			VocalDrone3d.aif	Stereo	1 or 2 then next							
WSQ			VocalDrone3e.aif	Stereo	longer use only at full park and when in no other zone or on way to Harvestworks							
WSQ			RhupTal BackingResonShortStereo	Stereo	infinite but time-based. ends when in a zone for more than 3 minutes (then it triggers Rl							
WSQ	NOT NEEDED ?		RhupTalTransition	Stereo	4							
WSQ			RhupTalFilterDelay16	Stereo	4							
J	Enter	no loops	Pigeons1.wav	Stereo	1							
~~J~~	~~trig random in list~~		~~Pigeons2.wav~~	~~Stereo~~	~~1~~	~~I'd like to use these, but trying to have less files. Will make a compc~~						
~~J~~	~~trig random in list~~		~~Pigeons3.wav~~	~~Stereo~~	~~1~~	~~I'd like to use these, but trying to have less files. Will make a compc~~						
J	trig random in list		Pigeons4.wav	Stereo	1	I'd like to use these, but trying to have less files. Will make a compc						
J	trig random in list		Pigeons5.wav	Stereo	1							
J	trig random in list		Pigeons6.wav	Stereo	1							
H	Enter	no loops	ForestBirdsDark4.wav	Stereo								
H	trig random in list		ForestBirdsDark1.wav	Stereo								
~~A~~	~~trig random in list~~		~~ForestBirdsDark2.wav~~	~~Stereo~~	~~COULD BE REMOVED IF MORE NEED TO BE TAKEN OUT~~							

Event List | Zone Key Info | QUESTIONS | new things to do 11-10

Figure 14.4 Section of the Event List for *Walkie Talkie Dream Angles* (Screenshot).

"song," were used as a way to keep listener-participants moving along, and as red thread throughout the piece. The tempo of the percussion was at a walking pace, and the repetition of various long sung phrases helped to connect the compositionally abstract electroacoustic pieces.

b. *Remember me?* The percussion and voice loops served to remind the listeners, especially those unaccustomed to abstract electroacoustics or soundscape compositions that what they were hearing was music. It was especially needed for audio work that could be mistaken for a similar overheard sound.

c. *How long has this been going on?* Whether using triggered audio files or algorithmic composition, it is good to make decisions regarding what happens if someone is engaging one part of the piece for longer periods of time. In WTDA, the amount of time participants remain within a single zone is measured. If it exceeds three minutes, a successive list of similar but increasingly simpler percussion loops are triggered, until the percussion nearly drops out altogether, along with the voice. When the participant continues their walk (crossing into new zones), the percussion and vocal sounds start up again. However, there is always at least a simple pulse playing, to assure the listener that the soundwalk is still active.

d. *When?* Tracking the time of day, the day of week, or even the time of year is yet another way to bring context or overturn expectations. Imagine a different composition at night, or in winter.

2. *Accounting for and using physical attributes of the environment*. The physical attributes of a particular landscape or architectural space will impact on AAR works. These include not only the *acoustics*, but also *traffic flow* of listeners (pedestrians or cars/bikes), *traffic density* and *what sort of activity is usually present*. All of these attributes can impact on what can be accomplished creatively.

a. *Acoustics*. Washington Arch is a defining visual image of the park. It was decided to work with the acoustics and reverb under the arch by recording some impulse response recordings and creating a convolution reverb with them. As a method for augmented reality, a very heavy dose of this locally sourced reverb was applied to loud "clacks" and noises recorded at the arch, to create electroacoustic pieces based on these sounds. Clips of these reverb-heavy sounds were organized so as to be triggered by any listener anywhere within 50 feet of the arch. Accentuating the reverb greatly expanded the area in which it is experienced. The liberally added reverb functioned as an audio suggestion of "archness," felt by listeners as soon as they approached that zone, in a subtle but reportedly effective way.

b. *As influencing traffic flow*. One should notice where people are walking—the paths or shortcuts ("desire paths") pedestrians take—to see which zones will be visited the most. In Washington Square Park, the central plaza and fountain are the heart of the park. Participants doing the WTDA soundwalk usually come back several times to the center of the park, which is open acoustically, and a giant echo chamber for any louder noises in the entire square. For this central area, a greater range of sounds was used (more audio file choices), to avoid too much repetition of material on the repeat visits. Also taken into account were the acoustics, and the ever-present cacophony of sounds from hawkers, musicians and other street performers, such that when overheard they would assist in creating the needed variety making each moment different from the next.

c. *Centers of activity influencing traffic density*. It is useful to consider where people congregate or make a lot of noise, or where walking might be slow because of traffic density. The fountain is a center of activity in the park. In the summer it's full of children and dogs and their minders cooling off, people yelling and the loud whoosh of water spraying. In the winter, the fountain is off and silent. Recordings made at the fountain were used to create electroacoustic vignettes. Some were quite realistic (all water and voices), and some very abstract—filtered as to leave only the rhythms and energy. When the water is off (in winter), the audio reminds listeners of the water's absence, even reportedly making some feel as if they might get wet. When the fountain is on in summer, its sound comingles with the composed vignettes.

3. *Use of environmental sounds*. Some decisions should also be made early in the process, about how to interact (or not) with environmental sounds, and about how these sounds might positively impact the composition and contribute to the AAR. All of the following are important to consider:

   a. *Balance*. What will the volume of potential overheard/real sounds be, especially in relation to the recorded/processed/generated sounds?
      - Should the outside sounds be blocked out? Incorporated? Take precedence?
      - What kind of headphones or listening devices will participants be using?
      - Is it the soundwalk solo or group activity? (Will participants interact with each other? How will it impact on the experience of the sounds to be heard?)

- Will the sounds heard need to be synchronized between participants in some way?
- Will participants be asked to use speakers as a way to introduce new sounds to the environment? How will those sounds interact with each other? Consider examples such as Phil Kline's *Unsilent Night* and previous work or interactive pieces with groups walking/bicycling with small speakers such as in Kaffe Matthews's *Sonic Bike* projects, and Lainie Fefferman and Jascha Narveson's *Gaits Soundscape* for Highline Park, NYC.[12]

b. *Mix*. The composer should consider what other aspects of the EQ, mix, effects or audio production might be influenced by his or her choices of balance and sound delivery.

4. *Accounting for natural sounds/animals*. Omnipresent in many outdoor AAR pieces might be sounds of birds, insects and other local animals. In the case of WTDA, a deliberate choice was made to interact with the natural soundscape of Washington Square Park, where the natural sounds are surprisingly dense in some areas and in others disturbingly absent.

a. *Adding and amplifying*. An "over the top" mix of birds and insect sounds was created for wherever it was felt their presence could be highlighted or given emphasis. If the composer felt these sounds were missing in a particular spot, she put them back (e.g. adding swarms of bees to underscore that there are many fewer bees than just a decade ago). As elsewhere, electroacoustic processed versions were included as well. (Insect sounds are also heard in *Florambula*, Barbara Weber's U-GRUVE piece for Conservatory Garden in New York City.)

b. *Inventing a fairytale forest*. For WTDA, it was found that the very large and older trees at the northwest corner of Washington Square park (including the 310-year-old Hangman's Elm) suggested an old German fairytale forest scene, especially at sunset. So sounds were added: owls, night birds, both as unprocessed collage and electroacoustic processed sound that maintains a spooky character of the place but is obviously electronic music. This evocative audio is there day and night, but at sunset and at night there are also the resident birds and crickets adding their own sounds to the mix.

c. *Animal gatherings*. Naphtali also noted that the pigeons congregate by day year-round by a statue in a little plaza in the west part of the park. So, her "flocking" pieces are there too, consisting of clouds of flapping wings and electroacoustic interpretations, to designate that entire area as belonging to those birds.

5. *Man-made sounds*. Urban soundscapes are an opportunity for interaction or exaggeration/commentary.

    a. *Construction sites*. For WTDA it was noticed that on the eastern perimeter of the park there had been continuous construction for several years. So, an electroacoustic construction zone was created from field recordings made on location. Serendipitous interactions ensue between the rhythms in the audio pieces and what a participant hears there when the construction site is active. Unpredictable and fun, these are the clearest and most musically accessible expression of aesthetics of WTDA as related to the aleatoric sound work of John Cage, the listening practices of Pauline Oliveros and others and what the composer hoped would open the participant's ears to new sounds and new ways of hearing.
    b. *Traffic/Street sounds*. Cars driving down fifth avenue must drive around Washington Square park to continue south. Because of this interruption, many cars careen impatiently around the northwest corner of the park to continue their trip downtown as quickly as possible. The juxtaposition of these rushing cars with the tranquility of the park and birds was an excellent opportunity for exaggeration and commentary with electroacoustic pieces based on drag races, raceway sounds and extreme automotive sounds using Doppler effect. At first these added sounds were nearly unnoticeable and needed to be made much louder and more exaggerated because, as discovered, the sounds of the actual cars going by were even more extreme than expected. When WTDA listener/participants realized what was going on, the piece fulfilled its original mission for the acoustic ecology festival—to highlight how much extra noise there is even in a "quiet" corner of this park. Listeners often had not realized how loud and intrusive these sounds were.

6. *What was not recorded*.

    a. *A word about musicians*. Naphtali avoided recording any street musicians in the making of this piece. This was because she is a musician and did not want to take advantage of any of their work, but also because their sounds are always in the park anyway—ever present and serendipitously included in the work, except late at night or in the cold of winter.
    b. *No firetrucks please*. A deliberate choice was made in WTDA to omit any firetruck or emergency response sounds from the piece for two reasons: (1) they are already ubiquitous and unlocatable in a big city and so always part of the general soundscape and (2) in the case of an actual emergency, using them in the walk could present a problem. Civically, the composer feels these sounds should not be

used as part of an outdoor piece in an urban environment and that they need to be reserved for their original purpose.

7. *Inaccessible places*. In three places inaccessible sounds were brought outside of their normal environment to be heard and considered by those who are excluded from them. Using accessibility as a determining factor for setting the perimeter of these zones, the physical attributes of the location were included as a means for recognizing and seeking to break down the social/political symbolism of exclusion and boundaries.

a. *Playgrounds* are only accessible to children and their caretakers (as they should be). They are a cornucopias of rich sounds, when heard up close, but from a distance (from outside) they are at best a rough din of children's voices. Audio pieces were created from these sources, and they were placed in a large area encompassing and extending the reach of two of the playgrounds, particularly accentuating the metal gates which are the literal gateway to this sound world.

b. *Library*. NYU's Bobst Library is also inaccessible to anyone without privileges there. Libraries are conventionally understood as very quiet spaces, but there are always noises present even there. To highlight both the inaccessibility of the space and its extreme, and supposedly quiet, atmosphere, pieces were created using highly amplified noises from recordings made inside the library. The inaccessible sound world of the interior of the library was brought outside the building onto the sidewalk, overlapping with the park slightly at the southeast corner. This audio can also be heard inside the library for anyone (with access) wanting some distraction from studying.

8. *Walking—travel vectors*. Expanding the idea of a zone as a trigger for an audio event, it is also useful to track the movement between zones, or use the direct mapping of GPS data as trajectories to control some aspect of that audio event or audio generation. This is similar to early computer-vision methods, which divided the field of vision into quadrants and tracked which quadrants were activated to determine direction and speed.

a. *Direction of travel*. Participants walking north/south from the park to a local gallery heard the sounds in a different sequence[13] simply by activating the zones in a different order. The direction of travel can also be directly tracked using the GPS to deliver completely different experiences, controlling any parameter of the sound, sound generation or mix. Tracking direction of travel could be especially useful for walks that are long but not wide.

b. *Speed of travel*. For Walkie Talkie Dream Garden (developed by the authors and launched in 2018 for the waterfront areas in north Brooklyn, NYC, and Hamburg, Germany), there are different experiences for walkers, runners and those who opt to hear the piece from the passing ferries out in the middle of the East River. As with direction of travel, speed of travel can be used to enhance the interactivity of the soundwalk by tracking zone activation, directly via the GPS, or simply by careful planning of sizing and placement of zones.

c. *Proximity (to a marker)*. Using detailed location information, it is possible to point a participant/listener to compose an interactive music score dedicated to a specific visual marker or artifact such as a sculpture or statue/artwork, in a small zone. Mixed media artist/researcher Adrian Hazzard excellently outlined his ideas regarding composing location-based soundtracks to create and resolve melodic and harmonic tension within four distinct stages of interaction: approach, arrival, engagement and departure, to create interactive and dramatic scores for works in a sculpture park.[14]

Another example from Barbara Weber's U-GRUVE piece *Orela* (written for New York's Lincoln Center Plaza) is that she successfully marries the walking of listeners around the iconic fountain to the cyclical reiteration of notes in a choral piece. Moving away from the fountain lowers the level of the drum patterns there. Filtered sounds and darkening timbres of sounds near the iconic fountain make it seem more intimate and change the mood. The ultimate effect is that the intensity increases as you get closer to the center, and the music is underscored by the white noise of the fountain enhancing the feeling of movement between various corners of the plaza and in relation to the fountain.

9. *Scene-setting versus underscoring*. With AAR, just as in film music, we make choices as to whether we are scene-setting and underscoring, creating literal or metaphorical sounds, to augment or replace the original sounds of our scene. Independent of the style of music being created, underscoring may be a compositional decision to replace the original soundtrack of a location (human or natural sounds) with a new one. But how would simply replacing the sounds with new ones be different than wearing headphones and listening to tracks of music as you walk around? The difference is that in AR the music has been consciously written for the place, can be more evocative than descriptive and can be adaptive to your movements. Just because one is walking near a fountain doesn't mean the composer must evoke water, or because the route of the walk goes near a statue of an angel that the composer must cue up choir sounds. That would be more akin to tone painting, and while that may be interesting to kids (maybe), it may not get to the heart of the psychogeographical experience of the place. It's a state of mind versus state of action. The primary difference between AAR and simply listening to tracks of music is the act of discovery driven by interactivity and indeterminacy.

10. *Historical and cultural context of location* can be a powerful way to engage the listener/participant in a composition, not simply by means of a voiceovered educational/cultural explanation as in a museum audio tour, but by mining hidden meanings/sounds of a particular location. Washington Square Park proved to be a treasure trove of material and inspiration. In some parts of the park, Naphtali was not inspired by the resident sounds, and in some cases she felt that to use them would be "Mickey Mousing" and too literal. She instead chose to connect to these locations in more symbolic, personal ways, mining stories/songs, memory, history and texts, hoping listeners could ponder the essence and themes of the park and location itself.

a. *The sounds of protest—historical and current*. The center and heart of the park is the plaza, where protests and public gatherings of all kinds, big and small, have been held going back at least 150 years. She used three different protest sound sources as material—two that she recorded on-site, and one from the 1960s. Given no file limitations, she would have included many more protests from different eras, but the three collages of protest sounds have proven to be sufficiently effective at causing listeners to recognize collective protest as a historic function of the central plaza of the park. If a listener passes through when a protest is in progress (which is often enough), they will hear these protest collages as augmenting the experience.

b. *Hidden stories—Marcel Duchamp*. In 1917, Gertrude Drick, Marcel Duchamp and four other artists snuck through an unguarded door to the spiral stairs inside of Washington Arch, climbed to the top and threw a party with balloons, toy pistols, lanterns and plenty of liquor, even a small fire to keep warm. The "Arch Conspirators" proceeded with pomp, to read aloud their proclamation, declaring independence for the "Free and Independent Republic of Washington Square." In homage in a vocal clip, for the Arch and Chess areas, Naphtali recorded herself singing the manifesto's opening "Whereas, Whereas, Whereas" the one word repeated, ad infinitum, as it was originally read, to create a reminder of the long history of Washington Square as a center of bohemian fun and rule-breaking.[15]

c. *Chess area—embedded texts, songs of decision and justice*. There have been competitive chess games going on in the southwest corner of the park day and night as long as the composer can remember. Dissatisfied with creating music from clocks and the sounds of chess pieces, she realized this zone should instead tie into a concept of justice and decision making using the symbolism of the chessboard as military and judicial. It was the best spot for the sung text for that zone: the George Washington quotation carved into Washington Arch, "LET US RAISE A STANDARD TO WHICH THE WISE AND HONEST CAN REPAIR." As Marcel Duchamp was also a chess player, the "Whereas, Whereas" audio clip plays here as well.

d. *Church bells*. Continuing on themes of social justice and action, on the south edge of the park across from the historic Judson Church a solemn electroacoustic composition using church bell sounds can be heard.

e. *Shakespearean voices*. In an area that is now a playground and inaccessible to any others than children and their caregivers, Naphtali put Shakespearean voices. These sounds are remembered by older denizens of park from street theater performances that used to be set there in the summer years ago. An older participant on a WTDA soundwalk once reminded the composer that left out of the piece were the voices of drug dealers selling their wares, voices/sounds that were pervasive in the park at that time.

11. *The role of indeterminacy*. The AAR experience can be engaging for a long period of time (or a larger area) by strategically designing the role of indeterminate sounds overheard from the environment and how people interact with the space. WTDA is different every time it is experienced for many reasons, all of them adding to the aleatoric nature of the work.

a. The sonic environment is constantly changing and only partially predictable. The choice was made to mix most of the work to maximize the interference of the outside sounds; the composer therefore prefers if the entire soundwalk is heard with earbuds or open headphones

b. Each participant chooses their own path and, therefore, which zones are activated.

c. Each participant chooses their own pace and, therefore, influences the percussion density (if they linger) and speed of changes of the sounds as they move through zones.

d. The zones in turn contain lists of audio files mostly playing in random order (so two people walking together may still get different versions of the piece). Limits can be set on how many times a file might be repeated.

e. No two soundwalk sessions will be the same, especially given other triggering variables such as time of day, weather, direction of travel (e.g. moving from north to south) and so on.

In the past twenty years, there has been an explosion in the number and types of projects using interactivity, which we encounter in every part of our lives as we work and play and use our smart devices and computers. Ideas about interactivity and indeterminacy for AAR relate directly to all this other work as well as larger-scale work by digital media artists in interactive performance systems and sound installations. Additionally, the literature and thinking on interactivity within the game music world, which has gotten considerably more sophisticated during this time, is vastly useful here.

14.5. Outro. (Conclusion)

AAR is a new medium, allowing the composer to create new site-specific work for many different kinds of spaces and experiences. We have only begun to scratch the surface of what is possible now that personal handheld devices have become so commonplace and GPS tracking have become such universal experiences.

This chapter has outlined many of the aspects and considerations necessary for constructing AAR systems and creating compositions for them. AAR represents the state of the art of interactive, location-based music, and suggests the evolutionary trajectory of our larger music-making experience.

It is a new channel for the public to acquire and enjoy music, as well as a new channel for composers to connect with their audiences. Ultimately, we are working to effect a large-scale shift in the public's view of what music can be, and how it can have emotional impact on our lives in the context of living, while at the same time transforming composers and audiences into collaborators, listeners into performers and environments into instruments.

14.6. Online Resources

Additional online resources related to this chapter can be found online at http://audioar.u-gruve.com/.

14.7. Recommended Projects and Readings

Please visit http:/u-gruve.com/projects-and-readings/ for the complete list.

14.8. Projects

Andreas Zimmermann and Andreas Lorenz
LISTEN: Contextualized Presentation for Audio-Augmented Environments (2003)
Virtual/contextual immersive museum guide via custom hardware
http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=970C577DE375C687412D9E70F7F4B2AF?doi=10.1.1.67.3987&rep=rep1&type=pdf

The LISTEN project was an early attempt to provide museum visitors with a contextually driven experience designed to "provide a personalized immersive augmented environment, an aim which goes beyond the guiding purpose. The visitors of the museum implicitly interact with the system because the audio presentation is adapted to the users' contexts (e.g. interests, preferences, motion, etc.), providing an intelligent audio-based environment."

================

Nigel Helyer and Dr. Daniel Woo
AudioNomad (2004)
Mobile application for Windows Pocket PC

A very early example of an AAR system, in which custom software is used to plot audio zones and GPS is used for tracking the participant's position. It also included head-orientation tracking for handling stereophonic panning. One implementation of the technology was for the 2006 presentation of *Syren*, a multichannel audio installation located in Sydney Harbour (Australia) in which listeners would ride on a boat through an array of 12 speakers, with the audio changing based on the boat's relative position to any given speaker.

================

Janet Cardiff
"Her Long Black Hair" (2005)
Guided Soundwalk with Custom Audio Kit (CD Player/Photographs)
https://phiffer.org/hlbh/
From the site: "Janet Cardiff's *Her Long Black Hair* is a 35-minute journey that begins at Central Park South and transforms an everyday stroll in the park into an absorbing psychological and physical experience."

"The walk echoes the visual world as well, using photographs to reflect upon the relationship between images and notions of possession, loss, history and beauty. Each person receives an audio kit that contains a CD player with headphones as well as a packet of photographs. As Cardiff's voice on the audio soundtrack guides listeners through the park, they are occasionally prompted to pull out and view one of the photographs. These images link the speaker and the listener within their shared physical surroundings of Central Park."

==================

Christina Kubisch
Electrical Walks (2004-present)
Soundwalk with customized wireless, electromagnetic-sensitive headphones
www.christinakubisch.de/en/works/electrical_walks

Kubisch has found a way to transduce the electromagnetic fields within urban environments through customized headphones, allowing listeners to walk through a given city and hear its electromagnetic emissions at various points along the way. Revealing many unexpected sources, each with their own frequency and rhythm, leads to wholly viable indeterminate musical compositions.

===================

Duncan Speakman
Various Works
http://duncanspeakman.net/locative

Speakman's output in the area of location-based music is prolific, with projects going back as far as 2006, in which he's explored and examined a number of formats for delivering locative sound/music experiences.

Here are a few recommended projects:

http://duncanspeakman.net/boundary-songs/
http://duncanspeakman.net/a-song-for-50-hearts/
http://duncanspeakman.net/for-every-step-you-take-i-take-a-thousand/

==================

Bluebrain (Hays Holladay and Ryan Holladay, composition and production)
Bradley Feldman (developer, creator of SSpace, the software used for the app)
Central Park: Listen to the Light (2011)
Mobile App for iOS
https://vimeo.com/channels/539003/29630558
www.nytimes.com/2011/12/08/arts/music/bluebrains-app-central-park-listen-to-the-light.html

Sometimes touted as "the world's first location-aware album," the Holladay brothers created a suite of original music for Central Park that was broken down into its parts and mapped to zones across the park. Using custom software developed by Bradley Feldman, the piece was delivered as an iOS app in 2011. The Holladays have also created similar experiences for the National Mall in Washington, DC, and R Park in Jackson, Wyoming.

=================

Josh Kopeček and Yoann Fauche
Echoes (2013-present)
Soundwalk generation software platform
https://echoes.xyz/

Echoes is a cloud-based platform for creating soundwalks. Though predominantly text-oriented, rich media files can be sound-art/musically oriented, though there is no temporal synchronization featured.

=================

Kaffe Matthews with Dave Griffiths et al.
Sonic Bike (various projects 2008-present, since 2013 as Bicrophonics Research Institute)
Custom fitted bicycles with audio speakers, GPS with custom hardware/software.
http://sonicbikes.net/sonic-bike/

The *Sonic Bike*'s onboard custom software allows for mapping of sounds onto multiple geographic areas/zones in cities and landscapes, to play content over the speakers "that changes depending on where the cyclist goes and how fast they ride." The composers (Matthews and her collaborators) have created a wide range of community-based experiences and location-driven musical compositions since 2008. Recent extensions of the project include the *Sensory Bike* (*Sonic Bike* as an instrument that can be played) and *Sonic Kayak*.

Notes

1. "The Recording Academy 'plays' New York City through AR leading up. . . ." 26 Jan. 2018, www.thedrum.com/news/2018/01/26/the-recording-academy-plays-new-york-city-through-ar-leading-up-grammys. Accessed 3 Feb. 2018.
2. "Evolution of Indoor Positioning Technologies: A Survey—Hindawi." 21 Feb. 2017, www.hindawi.com/journals/js/2017/2630413/. Accessed 3 Feb. 2018.
3. "GPS Beatmap on Vimeo." 2 Sep. 2009, https://vimeo.com/6402527. Accessed 3 Feb. 2018.

4. "Sonic Bike" http://sonicbikes.net/sonic-bike/. Accessed 15 June, 2018
5. http://sonicbikes.net/about-us/. Access 15 June, 2018
6. "Making Musical Apps—O'Reilly Media." http://shop.oreilly.com/product/0636920022503.do. Accessed 4 Feb. 2018.
7. "Getting Started with the Fluxamasynth Shield | Modern Device." 9 May. 2013, https://moderndevice.com/documentation/fluxamasynth-quickstart-guide/. Accessed 6 Feb. 2018.
8. "iOS, OSX and Android Audio SDK, Low Latency, Cross Platform, Free." http://superpowered.com/. Accessed 3 Feb. 2018.
9. New York City Department of Parks and Recreation, 2018. Washington Square Park [6 February 2018]. Available from: www.nycgovparks.org/parks/washington-square-park/
10. Defined as "playful, inventive strategies for exploring cities . . . just about anything that takes pedestrians off their predictable paths and jolts them into a new awareness of the urban landscape." *A New Way of Walking*, Joseph Hart, Utne Reader July 2004. www.utne.com/community/a-new-way-of-walking.
11. The Google "My Maps" feature allows for drawing and measurement of boundaries, and easy sharing of maps.
12. Friends of the High Line. The Gaits Soundscape 2017 In: The Highline Blog. 17 December 2017 [11 February 2018]. Available from: www.thehighline.org/activities/the-gaits-soundscape-2017
13. WTDA's first version of the walk for NoiseGate Festival extended into SoHo on two north/south streets.
14. Hazzard, A. 2014. *Principles for composing location based soundtracks* [5 February 2018]. Available from: https://adrianhazzard.com/research/principles-for-composing-location-based-soundtracks/
15. Axelson, E.P., 2007. The Free and Independent Republic of Washington Square (Part II). In: The Daily Plant. 24 January 2007 [viewed 3 February 2018]. Available from: www.nycgovparks.org/parks/washington-square-park/dailyplant/20026

Further Reading

Augoyard, J.F., Mccartney, A., Torgue, H. and Paquette, D., 2006. *Sonic Experience: A Guide to Everyday Sounds*. Montreal: McGill-Queen's University Press.

Filimowicz, M., 2015. The Mobile Augmented Soundscape: Defining an Emerged Genre. *Hz Journal*, no. 20. Viewed Feb 7, 2018 www.hz-journal.org/n20/filimowicz.html

Hart, J., 2004. A New Way of Walking. *Utne*, Ogden Publications, Inc, June 27, 2018. Viewed Feb 7, 2018 www.utne.com/community/a-new-way-of-walking

Hazzard, A., 2014. *Principles for Composing Location Based Soundtracks*. Viewed Feb 7, 2018 https://adrianhazzard.com/research/principles-for-composing-location-based-soundtracks/

Hazzard, A., Benford, S. and Burnett, G., 2014. You'll Never Walk Alone: Composing Location-Based Soundtracks. In: *Proceedings of the International Conference on New Interfaces for Musical Expression. NIME'14*, June 30–Jul 03, 2014. London: Goldsmiths, University of London, pp. 411–414.

Horowitz, S. and Looney, S.R., 2014. *The Essential Guide to Game Audio: The Theory and Practice of Sound for Games*. Burlington, MA: Focal Press.

Koutsomichalis, M., 2011. Site Specific Live Electronic Music: A Sound-Artist's Perspective. In: *Proceedings of the Electroacoustic Music Studies Conference, Sforzando!* June. New York. Viewed Feb 7, 2018 www.ems-network.org/IMG/pdf_EMS11_Koutsomichalis.pdf

Lafrance, A., Feb 19, 2016. Hearing the Lost Sounds of Antiquity. *The Atlantic*. Viewed Feb 7, 2018 www.theatlantic.com/technology/archive/2016/02/byzantine-angel-wings/470076/

Matthews, K. et al. Feb 12, 2018. *Bicrophonic Research Institute—Makes Music and Audio Landscapes to Be Triggered by You the Cyclist*. Bicrophonic Research Institute. Viewed Feb 7, 2018 sonicbikes.net/

McCartney, Andra, 2017. Soundwalking: creating moving environmental sound narratives. In: Gopinath, S and Stanyek, J, (Eds.), *The Oxford Handbook of Mobile Music Studies,* Volume 2. New York: Oxford University Press, pp. 212–237. https://soundwalkinginteractions.wordpress.com/2010/09/27/soundwalking-creating-moving-environmental-sound-narratives/

Oliveros, P., 2005. *Deep Listening: A Composer's Sound Practice*. Lincoln, NE: Deep Listening Publications.

Paquette, David and Andra McCartney, 2012. Soundwalking and the Bodily Exploration of Places. *Canadian Journal of Communication,* 37, no. 1, pp. 135–145. www.cjc-online.ca/index.php/journal/article/viewFile/2543/2286

Serafin, S., Erkut, C., Kojs, J., Nilsson, N.C. and Nordahl, R., 2016. Virtual Reality Musical Instruments: State of the Art, Design Principles, and Future Directions. *Computer Music Journal*, 40, no. 3, pp. 22–40. doi:10.1162/comj_a_00372

Urban Squares website. A "Collection and Analysis of Rediscovered Urban Space" with information about psychogeographical practices and "neighborhood portraits." http://urbansquares.com/07psychogeographyNEW.html

Westercamp, H., 2007. Soundwalking. In: Carlyle, A. (Ed.), *Autumn Leaves, Sound and the Environment in Artistic Practice*. Paris: Double Entendre, p. 49. Viewed Feb 7, 2018 www.sfu.ca/~westerka/writings%20page/articles%20pages/soundwalking.html