Remix_Dance Music 3: Improvisatory Sound Diffusion on Touch Screen-Based Interface

Jaeseong You
City University of New York Graduate Center
365 Fifth Avenue
New York, NY 10016
jaeseongyou@gmail.com

Red Wierenga
redspect@gmail.com
Arsid Ketjuntra
thearsid@gmail.com
Teerapat Parnmongkol
teerapatch88@gmail.com

ABSTRACT
Remix_Dance Music 3 is a quadraphonic quasi-fixed tape piece; an improviser can operate 60 pre-designed audio files using the Max/MSP-based interface. The audio files appear as icons, and the improviser can touch and drag them to four speakers in a miniature room on a tablet. Within the fixed duration of six minutes, the performer can freely activate/deactivate the audios and realize 32 different sound-diffusing motions, generating a sonic structure to one's liking out of the given network of musical possibilities. The interface is designed to invite an integral musical structuring in the dimensions of performatively underexplored (but still sonically viable) parameters that are largely based on MPEG-7 audio descriptors.

Keywords
Novel controllers, interface for musical expression, musical mapping strategy, music cognition, music perception, MPEG-7

1. INTRODUCTION
In designing new musical interfaces, the importance of mapping has been recognized as a musically creative process in and of itself [8]. Accordingly, there have been many intriguing studies on different ways one can map the interface inputs to the sonic outputs [4, 5, 6]. However, in regards to handling of concrete sounds, the musical mapping strategies have still remained relatively unexplored despite the already widespread popularity of DJ controllers and samplers [1].

The present project, instead of aiming for sophisticated mapping, allows the controller only the extremely simplified tasks of activation/deactivation of audios and their displacement. Even these two operations are always interconnected; if one wants to activate an audio file, then he has to make decision as to which sound-diffusion motion the audio is to take. The mapping strategy, therefore, is many-to-absolute-one, which might not be satisfying for some virtuosity-orientated players [5]. While the control option is admittedly and purposefully limited, the most sophisticated interaction between the player and the interface happens in the stage of handling of the information of the audio files. The interface is aimed to quickly order them by a variety of different criteria, and the performer is asked to give his real-time response as to how to interpret those criteria musically.

Based on Perry Cook's principles for designing new computer music controllers, the present project aims to make music, rather than an instrument or a controller [2]. *Remix_Dance Music 3* is designed to merge composition, performance, and instrument/interface into one cohesive musical process. The ultimate goal is to turn this musical process into six exciting minutes of improvisatory experience exciting for both the player and the audience.

NIME'13, May 27-30, 2013, KAIST, Daejeon, Korea.

2. DESIGNING THE AUDIO
The 60 precomposed audio files constitute the entire web of musical possibilities, in which the performer can freely navigate himself through. Thanks to the manageable number of the audios and their carefully crafted characters as musical lines of significant independence, the piece assumes a more tangible, consistent, and recognizable musical identity of its own despite its improvisatory nature.

The audio descriptors selected for the present project can be grouped into three categories: concrete domain, frequency domain, and time domain. All the parameters, with the single exception of 'instrumental name (please refer to 2.1.1),' can be quantified with a single floating-point value. The computation of such values should take place before the performance, so the audio files can be named in the format of list that takes the values of the audio descriptors as its components: (instrument name, audio descriptor 1, audio descriptor 2, …, audio descriptor n).

With these sonic descriptors appearing always in the same order in the list, the computer can perform quick real-time ordering of the audio files by simply looking at the 'n'th components in the lists and comparing their values.

2.1 Concrete Domain

2.1.1 Instrument Name
Instrument name is the single most powerful descriptor of the audios. Particularly given the musical context of *Remix_Dance Music 3*, in which the sound sources are predominantly from a variety of drum samples, one can get a surprisingly precise idea of how an audio in question would sound like given its instrument name. As this is the most useful asset for the improviser, the audio-icons will bear two-letter abbreviation of the different instrument names, and so the player can rely on this constant dimension anytime.

2.1.1.1 Percussion Lines
Out of the 60 audio files, 55 constitute the percussion lines, and they are classified into the following 11 instrument names: bass drum, whose abbreviation is bd, snare drum (sd), hihat_closed

(hc), hihat_open (ho), cymbal (cy), click (ck), clap (cp), shaker (sk), cowbell (cb), woodblock (wb), and tom (tm).

2.1.1.2 Processed Noises

5 audio files contain steady processed noise, respectively named from noise_1 (n1) to noise_5 (n5).

2.1.2 Pattern Periodicity

Each and every precomposed audio is made of regularly recurring gestures except the five processed noises. Therefore, their periodicity can be a particularly useful piece of information in manipulating the macro-level sensation of regular pulse and tempo. Therefore, the periodicity is employed as one of the parameters, expressed in the unit of millisecond.

2.2 Frequency Domain

All the audio descriptors under this section (2.2) and the next (2.3) are selectively and exclusively borrowed from Tae Hong Park's Ph.D. thesis on timbre recognition [9]. The values are to be computed by his MATLAB-based software, ElectroAcoustic MuSic AnalYsis toolbox (EASY toolbox) [10]. As briefly mentioned, each value can be represented by a single floating point. Because each parameter is never judged in conjunction with other parameters, we can simply compare the relative values between the audio files within one parameter at a time, avoiding any complicated and error-inducing normalization process.

Some of the sonic parameters might seem arbitrary and excessively abstract as they are devised originally for machine learning. It is admittedly true that there is no guarantee that human can effectively hear what computer can hear. However, for that very reason, the parameters are selectively chosen based on the first author's listening. Also, many of the parameters belong to MPEG-7 audio descriptors that are widely used to describe varying contents of sound. With these parameters given, the judgment of which ones are more musically useful is entirely left to the improviser.

2.2.1 Net Inharmonicity

Net inharmonicity is defined as the sum of the errors between measured harmonics and their theoretical ideal harmonics (multiples of the fundamental frequency).

2.2.2 Harmonic Expansion/Compression

Harmonic expansion/compression literally shows how much the harmonic series is expanding or compressing.

2.2.3 Spectral Centroid

Spectral centroid is defined as the weighted mean of the frequencies present in a signal, which shows where the center of mass (center of gravity) of the spectrum of the given signal is.

2.2.4 Spectral Irregularity

Spectral irregularity computes the jaggedness/smoothness of the spectral envelop.

2.2.5 Spectral Flux

Spectral flux calculates the amount of frame-to-frame fluctuation of the given signal in time, measuring how quickly the power spectrum of the signal changes.

2.2.6 Log Spectral Spread

Log Spectral Spread is defined as the energy found between boundary pairs of frequency components.

2.2.7 Roll-off

Roll-off is defined as the frequency boundary where 85% of the total power spectrum energy resides.

2.2.8 Spectral Flatness

Spectral flatness is defined as the ratio of the geometric mean of the power spectrum to the arithmetic mean of the power spectrum.

2.3 Time Domain

2.3.1 Log-rise Time (LRT)

LRT refers to the logarithmic value of the time between a start-point and the maximum amplitude-point.

2.3.2 Temporal Centroid

Temproal centroid is defined as the energy weighted mean of the time signal.

2.3.3 Fundamental Frequency

Although the precomposed audios are primarily "pitchless" percussive sounds, the fundamental frequency is still employed as a valid sonic descriptor as the weak tonality of signal does not equate to no fundamental oscillation.

2.3.4 Zero-Crossing Rate (ZCR)

ZCR is defined as the rate of the time-domain zero crossings within a windowed portion of a sound normalized by the window length.

3. STURUCTURING THE INTERFACE

It might be somewhat difficult to precisely locate the present project within the context of the recent development of new instruments. As it deals exclusively with prerecorded audios, it is closer to DJ controllers [1]. However, while latest DJ controllers usually provide GUI for mixing and playing the samples and knobs/sliders for controlling auxiliary parameters [7], the present controller is more concerned with simplifying the input/output mapping to the extreme; the only control it offers is ultimately the binary amplitude option (of either silencing the audio file or playing it at its original amplitude), always coupled with the sound displacement choice.

To exploit the maximal potentials of the radically reduced control, the interface employs touch screen as its hardware, which has been praised for its immediacy, intuitivity, and agility [3]. The core-engine of the interface is developed in Max/MSP for its advantage in GUI and real-time play [8], and it is connected to iPad by means of Air Display application, which turns iPad into the secondary monitor of the laptop.

Another noteworthy feature of the present interface is the combination of sound displacement and the touch screen tactics. There have been a number of attempts in developing performative sound diffusion system based on the motion capture strategies, and they have proven to be visually and sonically quite intriguing. However, one possible drawback is that the player's engagement with the sound is destined to be first-person experience, which is not always desirable for improvisatory setting. In such systems, it is difficult for the performer to pull himself from the sound he is making and objectify the musical situation. Improvising over a certain length of duration with no pre-planned structure, he can easily fall into the myopic self-repetition that can be musically uninteresting.

With the present controller, the performer can map the precomposed audio icons over the virtual miniature room, in which the quadraphonic playback system is represented with simple and straightforward geometric shapes from a bird-eye view. The interface helps the improviser to envision himself out of the first-person contact with the sound of physical Now, so he can approach the improvisation more from a perspective of real-time composition.

3.1 The Graphic Surface

The following image is the diagram of the present interface. The precomposed audio files appear as small icons, and the performer can move them to 33 different regions in the virtual miniature room.

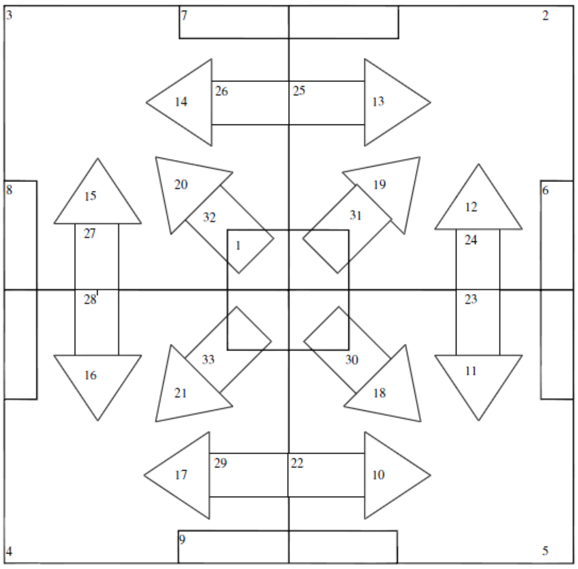

Figure 1. Controller Diagram

The central square (1) is called "neutral zone" where all the icons are initially located. Its function is two-fold: (a) it orders the icons by different parameters as the preset buttons are located in this area, and (b) the icons in this region do not sound (but they are perpetually playing as all the audios have to end simultaneously in 6 minutes), and so, by dragging them in out of this region, one can "activate/deactivate" the sound.

In the region (2), the icons sound only through the first quadrant speaker. The regions (3), (4), and (5) serve the same purpose with their corresponding speakers. In the region (6), the icons sound through the first and fourth quadrant speakers in equal distribution. (7), (8), and (9) also work similarly with their corresponding speakers. In the region (10), the icons sound first at the third quadrant speaker, and they QUICKLY move to the fourth quadrant speaker to stay there. The default value for this quick sound disposition is 4 seconds, which can be varied by the player within the range between 2 seconds and 8 seconds. The regions from (11) to (21) also play the role of the quick sound disposition with their corresponding speakers. In the region (22), the icons sound first at the third quadrant speaker, and they SLOWLY move to the fourth quadrant speaker within the default value of 10 seconds, which can be varied between 9 seconds and 20 seconds. The regions from (23) to (33) operate as the same slow sound diffusion with their corresponding speakers.

4. PLANNING THE PERFORMANCE

The improviser will be encountering two stages of performance: (1) grouping of the pre-constructed audio files by offered parameters and (2) activating/deactivating and displacing of them through the quadraphonic. In choosing what audio files to use, the performer is free to translate the abstract parameters back to his ordinary terminologies such as noisiness (ZCR), smoothness (spectral irregularity), richness (spectral spread), and so on. The maximal number of simultaneously sounding audios is recommended to be limited to 25 to 30, and so the density does not exceed the musically interesting limit. This also helps the performer to deal with only a controllable amount of sounds at one time.

The performance requires a quadraphonic setup with each speaker located nearby each corner of square room, angled slightly towards the center of the room where the audiences are seated. The performer should be in the place where he can match the location of the actual speakers as the way the speakers are represented on the controller.

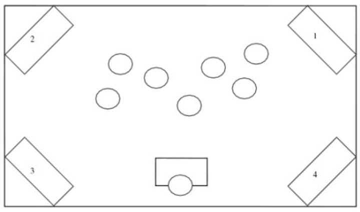

Figure 2. Stage Setup Diagram

5. RELATED WORKS

The followings are the previous works in which the first author has experimented with the similar musical approach.

5.1 Dance Music 1

Premiered at SEAMUS 2013, *Dance Music 1* is built with a similar character of material as *Remix_Dance Music 3*. This piece also shares the same approach of controlling the activation and deactivation of simultaneously running precomposed musical lines. The audio can be found under the following link:
http://ciociosanandbutterflies.bandcamp.com/track/dance-music-1

5.2 Dance Music 3

Dance Music 3 is a demo construction of *Remix_Dance Music 3*. It is a down-sized version of the present project in terms of both the duration and the number of precomposed audio files. *Remix_Dance Music 3* employs exactly the same collection of samples as *Dance Music 3*. The audio can be found under the following link:
http://ciociosanandbutterflies.bandcamp.com/track/dance-music-3

6. COMPROMISES

Park's EASY Toolbox project is still in the stage of development at this point. Consequently, the sonic description part of the present project has not been realized; for the purpose of performance, the improviser should be familiar with the nature all the audio files in advance. Also, some details of audio files differ from the previous description due to the minor changes introduced for aesthetic and practical reasons.

7. CONCLUSION

The goal of the present project is to create a musically viable improvisatory environment, seamlessly integrative of composition, performance and instrument/interface. The performer can operate the controller by displacing the sound icons over the graphic surface to activate/deactivate and diffuse them. The controller leverages the already-proven intuitivity and agility of the touch screen-based interface. Additionally, it helps the improviser to easily narrate the sense of larger musical structure by controlling the sonic situation from the bird-eye view; despite the real-time setting, the player can take advantage of a stronger "composerly" power. Also, the project heavily relies on ordinarily underrepresented sonic parameters, inviting unconventional musical structuring tactics. Finally, the previous related works, *Dance Music 1* and *Dance Music 3*, constructed with the similar materials based on the same methodologies, support and prove the aesthetic validity of the present project.

8. BIOGRAPHY

8.1 Jaeseong You

Jaeseong You is a New York City-based composer. Finishing his B.A. at New York University, he is currently an Enhanced Chancellor's Fellow at City University of New York Graduate Center in pursuit of his PhD, and he will be moving to NYU Steinhardt to study with Tae Hong Park. His academic research interests lie in quantification of perceived amplitude and audio-based music analysis.

Having studied with David Olan, Elizabeth Hoffman, Richard Carrick, Douglas Geers, and Arthur Kampela, You actively composes both electronic and acoustic music. His acoustic works have been performed by ensembles like MIVOS Quartet, Talea Ensemble, Cygnus Ensemble, and American Modern Ensemble among many others. As an active member of the computer music scene, he reviewed works for ICMC 2013, and his electronic works have been invited to SEAMUS 2013, IEMF 2013, and NIME 2013.

8.2 Red Wierenga

Red Wierenga is a pianist, keyboardist, respectronicist, improviser, and composer currently based in New York City. His longest creative association is with the Respect Sextet, called "a group which has released one of the most compelling recordings of the year," by the Wall Street Journal, and "one of the best and most ambitious new ensembles in jazz" by Signal To Noise.

He performs and records in a wide array of musical settings, from free improvisation, jazz and new music to rock, pop, and world musics. He has performed and/or recorded with artists including the Respect Sextet, the David Crowell Ensemble, Ensemble Signal, Brad Lubman, Salo, the Fireworks Ensemble, The Claudia Quintet, and others.

8.3 Arsid Ketjuntra

Arsid Ketjuntra was born in Bangkok in 1983. He began his music education at the College of Music, Mahidol University. As a music technologist, Ketjuntra is interested in digital sound synthesis and recording technology. He took private composition lessons from Surat Kemaleelakul during his undergraduate study, which continued after graduation. Having finished his two-semester lectures on Music Theory at the Mahidol University, he moved to Brooklyn, New York, and has been studying music composition with Aya Nishina and Tania León. He is also studying computer music at the Brooklyn College Center for Computer Music with Douglas Geers.

8.4 Teerapat Parnmongkol

Teerapat Parnmongkol (Lemur Onkyokei) is sound artist and improviser from Bangkok, Thailand, and currently based in New York City. Teerapat studied music technology and contemporary music at ChandraKasem University, and he has graduated from SAE with diploma of Audio Engineer.

Parnmongkol started off as a freelance composer and quickly found himself interested in minimalist approach and perceptive time collapse. As a result, he began his individual side project of Lemur Onkyokei extensively employing ethereal sounds and experimental noises. Most of his works are focused on the repetitive, electro-acoustic process, interactive media, and cognitive dissonance and consonance. Parnmongkol co-founded GAD and Sonic Scape Orchestra and remains as an active member.

9. REFERENCES

[1] Anderson. T. H. 2003. "Mixxx: Towards Novel DJ Interfaces" In *Proceedings of the Conference on New Interfaces for Musical Expression*.

[2] Cook. P. 2009. "Re-Designing Pirnciples for Computer Music Controllers." In *Proceedings of the Conference on New Interfaces for Musical Expression*.

[3] Geiger. G. 2006. "Using the Touch Screen as a Controller for Portable Computer Music Instruments." In *Proceedings of the Conference on New Interfaces for Musical Expression*.

[4] Hunt. A. 1999. Radical user interfaces for real-time musical control. Ph.D. dissertation, York University.

[5] Hunt. A. and M. Wanderley. 2002. "Mapping performer parameters to synthesis engines." Organised sound, 7(2):97-108.

[6] Hunt. A. and M. Wanderley, M. Paradis. 2003. "The Importance of Parameter Mapping in Electronic Instrument Design." Journal of New Music Research, 32(4): 429-440

[7] Lippit. T.M. 2006. "Turntable Music in the Digital Era: Designing Alternatives Tools for New Turntable Expression." In *Proceedings of the Conference on New Interfaces for Musical Expression*.

[8] O'Sullivan. Liam. and D. Furlong, F. Boland. 2012. "Introducing CrossMapper: Another Tool for Mapping Musical Control Parameters" In *Proceedings of the Conference on New Interfaces for Musical Expression*.

[9] Park, T. H. 2004. Toward Automatic Music Instrument Timbre Recognition. Ph.D. dissertation, Princeton University.

[10] Park, T. H. 2009. "EASY Does it." In *Proceedings of The International Society for Music Information Retrieval*.

[11] Rovan, J. B. and M. Wanderley, S. Dubnov, P. Depalle. 1997. "Instrumental gestural mapping strategies as expressivity determinants in computer music performance." In *Kansei, The Technology of Emotion, Proceedings of the AIMI International Workshop*. Genoa: Associazione di Informatica Musicale Italiana. 68-73.