Journal of Smart Technology Applications
Vol.1, No.1 (2020), pp.21-28
http://dx.doi.org/10.21742/jsta.2020.1.1.04

Exploration of Practical Information Entry Technique of One-Shot Drum Audio Samples

Chul-Hee Lee[1] and Sean Han[2]

[1]*Kyunghee University, South Korea*
[2]*Valor International Scholars, South Korea*
[1]*beat344@hanmail.net*, [2]*hsh4jesus@gmail.com*

Abstract

This research gives a compelling strategy on data section of one-shot drum audio samples with a high recurrence of utilization in music creation. One-shot drum audio samples have huge preferences of being utilized quickly in music creation. Be that as it may, users must bear the burden of listening to each and every sample sound source in light of the fact that the data entry method for them has not been created. Accordingly, this study centers around the five most significant sorts of data in the one-shot drum audio sample: length, peak energy frequency, overtones, tapping sounds and transient shapes. This paper uncovered the impacts of these components on sound source and tone. Also, by introducing objective estimation criteria and techniques for evaluating the samples, it was conceivable to make systematic data entry. This provided a more open approach to users to productively oversee and use the samples by entering the figures for each identified component in the media data of the sample music file properties. We hope this study will reduce the working hours of music production drastically, and we hope that the producers who produce audio samples will also use it actively.

Keywords: *Drum sample, Audio sample, One-shot sample, Music production, Beat making*

1. Introduction

Early drum machines could only use drum sounds by combining white noise and sine waves, but with the release of Linn's LM-1 in 1980, it was possible to sample and program actual drum audio sources [1]. Since then, with the release of relatively inexpensive digital samplers such as E-mu's SP1200 and Akai's MPC, music production using sampling techniques has rapidly increased. As the sampling technique became more common, Ratcliffe proposed a classification method for dividing the sound sources used in sampled music into short/isolated fragments, loops and phrases, large elements, and transformed materials according to their length [2]. This study deals with the sound sources of shortly divided pieces that correspond to drum one-shot audio samples.

The starting point of the popularization of sampling techniques can be seen as hip-hop music in the late 1980s. Hip-hop producers at the time used digital samplers to create music with sophisticated sound source editing techniques [3]. At this time, due to the introduction of digital samplers, such as SP1200 and MPC, loops, phrases, and short/isolated fragments sampling techniques could be used easily, resulting in numerous beats in the golden age of

Article history:
Received (May 13, 2020), Review Result (June 25, 2020), Accepted (August 3, 2020)

hip-hop [4]. Also, with the appearance of audio sample CDs, one-shot audio samples that can be used immediately without a separate editing process have been popularized. Currently, due to the development of the Internet, there are various websites that provide audio samples. In particular, Splice, a website launched in 2013, offers millions of samples, loops, FX, and presets, which are used by millions of musicians [5]. However, the currently used audio sample information writing method has to be improved because it is impossible to predict the acoustic information of the sound source without a listening process. Therefore, this study aims to contribute to improving the efficiency of work by suggesting a method of effectively displaying acoustic information on drum one-shot audio samples, which are becoming more popular in music production as such.

2. Problems of the existing information writing method for drum one-shot audio samples

Audio sample CDs are generally classified by specific genre or style of music. The production company selects a specific genre and style that can be commercialized, then produces and releases audio samples corresponding to them, and users listen to an audio sample CD corresponding to the music they want to make, and then select a specific sample. When handling the sample, the genre and style entered by the production company have the greatest influence on the selection of audio samples. Users typically keep the title of the album entered by the production company even if the audio sample CD's sound source is moved to a storage device in the form of a file. This classification method is commonly used because of the advantage of being able to predict the genre and style that audio samples are aimed at, but has a limitation in that it is difficult to predict the acoustic characteristics of each sample.

On the other hand, websites such as Splice provide simplified images of the waveform of the sound source along with tags for the type of audio sample, instrument, and genre. This method helps predict an approximate envelope through visual information on the audio sample waveform. However, information that determines the tone such as pitch and overtone cannot be grasped. In this way, in the process of finding the desired drum one-shot audio sample, the user spends an excessive amount of time listening to the numerous audio samples he has. This can be regarded as a problem due to the limitation that information of the audio sample cannot be effectively written. Therefore, it is expected that the efficiency of the working speed can be greatly improved if the information writing method of the drum one-shot audio sample can be effectively improved.

3. How to write effective information of drum one-shot audio samples

3.1. Key information of the drum one-shot audio sample source

Acoustic information of sound can be subdivided countlessly according to purpose and classification, but among a variety of information, five types of information that have a major influence on the acoustic characteristics of the drum one-shot audio sample are selected as follows, and the analysis method of each information was presented.

3.1.1. Length

First, the duration of the sound source must be considered when writing information about drum one-shot audio samples. In particular, in the case of the hi-hat cymbals, the closed hi-hat

and the open hi-hat have similar elements such as pitch and overtone, but the difference in duration is large as shown in [Figure 1] below.

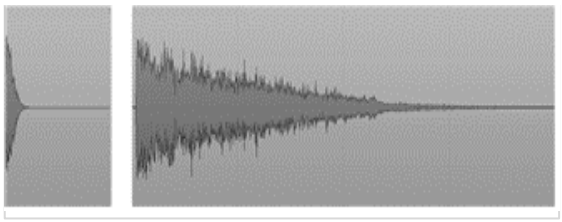

Figure 1. Waveform of closed hi-hat(left) and open hi-hat(right)

Note that the duration of the sample may differ from the length of the entire sound source. If you look at the waveform of the left closed hi-hat in [Figure 3], you can see that the length of the sound source is longer than the actual duration of the audio sample because there is a gap after the sample is finished. Therefore, since it is difficult to determine the actual duration of an audio sample using only the total length information of the sound source, it is very important to measure the actual duration and provide information. If the duration of the audio sample is measured and information is provided in a numerical value, the user will be able to take the first step towards selecting a sample even if he or she does not listen to the sample. By using the principle of a threshold level in dynamic effectors, like a gate and compressor, it is possible to measure the duration of an audio sample.

3.1.2. Peak energy frequency

Among the main factors that determine the tone of a drum sound, the peak energy frequency depends on invariant factors such as the size and thickness of the instrument and variable factors such as tuning. Hyung-Jun Lim analyzed the peak energy frequency for each genre per year for the bass drum in his paper as shown in [Table 1] below.

Table 1. Peak energy frequency by year for each genre of the bass drum [6]

	Pop	Dance	Rap	Rock	R&B
Year 2000	60	70	85	55	58
Year 2003	63	60	85	85	63
Year 2006	50	50	75	105	45
Year 2009	90	48	100	60	55
Year 2012	90	45	180	80	45
Year 2015	80	95	60	63	95
Average	72.2	61.3	97.5	74.7	60.2

**Unit : Hz*

Of course, it is not possible to generalize drum audio samples aimed at a specific genre because the number of sample sources analyzed in his thesis is insufficient, but it is possible to confirm that there may be differences in peak energy frequencies for each sound source. Therefore, by detecting the peak energy frequency of an audio sample and providing numerical information, users can select audio samples similar to their desired sound source and exclude the rest from the listening process, significantly improving the speed of their

work. The method of detecting the peak energy frequency is possible with a peak frequency measurement using a frequency analyzer.

3.1.3. Over tone

Unlike instruments that maintain a constant tone, such as a piano, a guitar, string instruments, and wind instruments, drums tend to have a frequency that exceeds fundamental tone in an overtone structure rather than a harmonic structure. Even for the same instrument with similar fundamental frequency, the tone varies greatly depending on the structure of the overtone. When the high-frequency overtone sound pressure is high the sound gets brighter and clearer. While, the sound gets darker and softer if the high-frequency overtone sound pressure is low.

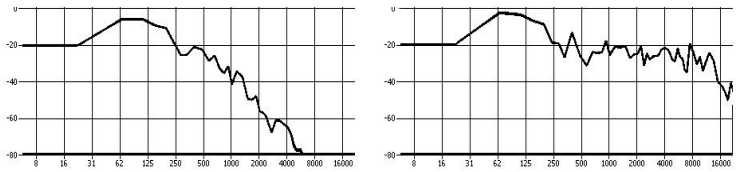

Figure 2. Audio spectrum of two different bass drums

The preceding [Figure 2] shows the audio spectrum of two bass drum audio samples with similar fundamental frequency, but significantly different overtone structures. The sample on the left has a low high-frequency overtone sound pressure and the sample on the right has a high high-frequency overtone sound pressure. Therefore, if the sound pressure of the high-frequency overtone of the audio sample is detected and information is provided in a numerical value, the user can preferentially select audio samples close to the desired tone, thereby greatly saving the time spent in the listening process. In order to quantify the overtone, it can be calculated as the difference in sound pressure for each frequency band detected in the spectrogram.

3.1.4. Tapping sound

Drum instruments such as bass drums, snare drums, and tom-tom drums show a momentary shift of pitch from high notes to low notes when playing. This phenomenon occurs due to the process of rapidly returning to the original sound by pressing the head at the moment of the stroke. It is a common phenomenon found in membranophones that are played by covering a drum barrel made of wood or metal with a head [7].

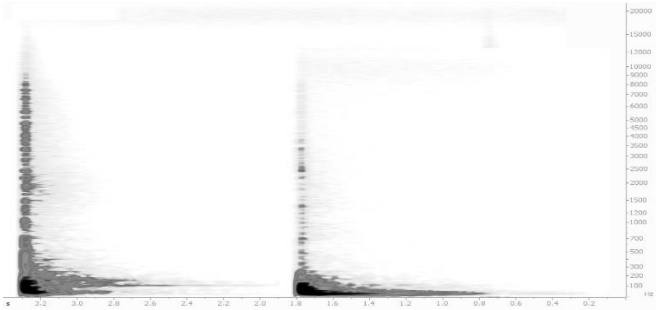

Figure 3. Audio spectrogram of different playing intensities of the same bass drums

[Figure 3] above is a spectrogram of playing different intensities with the same bass drum. The left side is the strongly played sound source, while the right side is the weakly played sound source, and the sound pressure of the weakly played right sound source was amplified to be the same as the left sound source in order to set the sound pressure of the two sound sources to be equal. Even at the same sound pressure, it can be seen that the pitch variation of the strongly played left sound source is significantly larger and clearer than the weakly played right sound source. Therefore, if you provide information in numerical terms after detecting the degree of pitch change of the audio sample, you can actively refer to it when selecting the desired tone sample. The degree of tapping sound of the sample can be quantified based on the width of the pitch change in the spectrogram.

3.1.5. Transient shape

One of the most important plug-ins used in the recent popular music mixing work is the devices related to the transient shape. In the drum audio sample, the transient shape can be divided into a shape with an emphasis on attack and a shape with an emphasis on sustain. In the form where the attack is emphasized, the sound pressure sharply decreases after the peak level, so that a strong tapping sound can be felt, and the form in which the sustain is emphasized amplifies the sustain level after the peak level, resulting in a rich sound.

Therefore, if information about whether the transient shape is close to the attack emphasis shape or the sustain emphasis shape is provided numerically, it will be useful for audio sample selection. Transient shape measurements can be quantified through changes in sound pressure over time detected in the spectrogram.

3.2. To enter information for a drum one-shot audio sample

The simplest way to enter information analyzed through the preceding process into an audio sample is to write it in the file name. However, there is a fatal disadvantage in this case that it is difficult to arrange sequentially into five key information categories. Therefore, for effective information entry, the following methods are proposed.

3.2.1. Complete file properties for existing Windows operating system

The five methods of recording sound source information presented prior to the drum one-shot audio sample are possible by entering each of the media information in the file properties of the existing Windows operating system into "Participatory Musician, Album, Year, #, Genre." It provides information for participating musicians by recording duration in ms(milliseconds) units, peak energy frequency in the album in Hz(hertz) units, difference in over-toned sound pressure to peak energy frequency sound pressure in dB(decibel) units for the year, # in st(semi-tone) units for tapping sound, and transient shape in % units for the genre.

이름	참여 음악가	장르	앨범	#	연도
attack-kick01.wav	200ms	95Hz	-20dB	24st	30%
attack-kick02.wav	210ms	100Hz	-18dB	22st	35%
attack-kick03.wav	220ms	105Hz	-16dB	20st	40%
attack-kick04.wav	230ms	110Hz	-14dB	18st	45%
attack-kick05.wav	240ms	115Hz	-12dB	16st	50%
attack-kick06.wav	250ms	120Hz	-10dB	14st	55%
attack-kick07.wav	260ms	125Hz	-8dB	12st	60%
attack-kick08.wav	270ms	130Hz	-6dB	10st	65%

Figure 4. Method of using information about audio samples

[Figure 4] above shows an example of the samples arranged in File Explorer based on the information entered in the audio sample. The advantage is that the samples can be sorted by similar sound source information, effectively reducing the time to select the desired sound source, and objective and standardized information can be entered without the use of unstructured information methods such as arbitrary or abstract file names previously used. However, while the above method has the convenience of storing audio information with the general ability to use a computer, there exists a disadvantage that the name of the item and the actual information do not match, which could cause confusion. It is necessary to modify each item name to a name that conforms to the information of the sound source.

3.2.2. Modify item names using Python

As a way to improve the inconsistency problem between the items of media information and the entered information, we propose a coding technique through the use of software. Using the Python programming language (hereinafter referred to as Python) developed by computer programmer Guido van Rossum, audio information written in a sound source can be converted to be displayed as an item corresponding to the information. Python is the most popular and most accessible programming language in the world. The advantages of Python are described as extensive libraries, improved productivity, and its characteristic of being free and open source [8]. We chose Python because of its variety of libraries, and the fact that Python is open source and free provided more possibility to this idea. After extracting the information contained in the music file and storing it in variable values, the information name and information can be listed together to extract proper information and information values. Through this, musical information such as the duration of the sound source recorded in the file and the peak energy frequency can be displayed in items that match the corresponding information, such as “continuing time” and “peak energy frequency” rather than items that are not related to the information, such as “participating musicians” and “albums.” To implement such programming, eyeD3 [9] of Travis Shirk (hereinafter, Travis) should be used.

4. Conclusion

The drum one-shot audio sample is a file that has completed several steps to digitize the sound of an instrument. The frequency of use has increased significantly due to its advantage of immediate use when producing music. However, because the systematic method of recording the acoustic information of the sample was not developed, users had to listen to numerous audio samples one by one and find the file they wanted. Therefore, this study proposed a method that can effectively write the acoustic information of the drum one-shot audio sample to help the work become efficient.

First, in this study, in order to effectively write the information of the drum one-shot audio sample, five categories were selected for the main information of the sound source. The corresponding categories are duration, peak energy frequency, overtones, tapping sounds, and transient shapes. The duration is quantified by measuring the duration of the actual audio sample, which may differ from the length of the sound source, and the peak energy frequency was quantified by detecting the frequency corresponding to the largest sound pressure in the sample. Overtones depend on the material of the instrument or the audio plug-in applied and are the main determinants of the instrument's tone. The method to quantify this was not the absolute sound pressure, but an objective method that indicates the difference between the peak energy frequency sound pressure and the overtone sound pressure. The sound of tapping is produced by the change in pitch that is common due to the characteristics of the musical

instrument, and the degree of the tapping sound is indicated by measuring the width of the pitch that rises and falls rapidly in the early stages of the sample sound source. In the case of the transient shape, the attack emphasis form and sustain emphasis form were quantified.

Five pieces of information quantified by the above method could be written in the file properties of the existing Windows operating system. In the properties of the audio sample file, among media information, information was provided with the duration in ms for the participating musician, the peak energy frequency in Hz for the album, the peak energy frequency in the year, the difference in the degree of reduction of the overtone sound pressure compared to the sound pressure in dB, the pitch change width corresponding to the sound of the strike sound in st, and the genre in % units. By referring to this information, the user can roughly predict the acoustic characteristics of the sample sound source without a listening process, and can effectively manage the audio samples by sorting each element in the explorer. In addition, the name of the item that does not match the information of the sound source, which is a problem that usually occurs at this time, was solved using Python programming. By using a program coding technique that inputs a sound source file through EyeD3 and stores five information values in a variable and outputs it as a list, confusion and inconvenience for users was prevented.

References

[1] https://www.theguardian.com/music/2009/jun/22/linn-oberheim-drum-machines, July 6, **(2020)**

[2] R. Ratcliffe, "A proposed typology of sampled material within electronic dance music," Journal of Electronic Dance Music Culture, vol.6, no13, pp.99-112, **(2014)** DOI:10.12801/1947-5403.2014.03.01.06

[3] J. R. Park, "An analysis method and applications of sound source sampling technique in popular music," Korea Journal of Popular Music, vol.-, no.19, pp.29-54, **(2017)**

[4] P. Edwards, "Concise guide to hip-hop music," St. Martin's Publishing Group, New York, pp.224, **(2015)**

[5] https://splice.com/features/sounds, July 6, **(2020)**

[6] H. J. Lim, "Frequency range and level analyze of kick drumin the representative music style genre," M.S. thesis, Department of Contemporary Media Music, Graduate school of Chugye University for the arts, Seoul, Korea, **(2016)**

[7] C. H. Lee, "A study on the characteristics of 808 synth bass for the k-pop music production education," Korea Journal of Arts Education, vol.16, no.1, pp.93-110, **(2018)**

[8] K. P. Naveen Reddy, Y. Geyavalli, D. Sujani, and S. M. Rajesh, "Comparison of programming languages: review," International Journal of Computer Science & Communication, vol.9, no.2, pp.113-122, **(2018)**

[9] https://eyed3.readthedocs.io/en/latest/index.html, August 5, **(2020)**

This page is empty by intention.