Exploring the Perceptual Relevance of Inherent Variability of Drum Sounds

Matthias Rath and Marcel Wältermann

Berlin University of Technology
Deutsche Telekom Laboratories
Quality and Usability Lab
`matthias.rath@telekom.de, marcel.waeltermann@telekom.de`
`http://www.qu.t-labs.tu-berlin.de`

Abstract. It is common "beat programming" practice to construct rhythmic sequences in a one-sample-per-instrument manner. Sounds of separate strokes on a drum, however, generally are not really perceptually identical, even when played with identical intention on the side of the player. Moreover, such unavoidable individual variations go beyond the effects of amplitude scaling or simple linear (e.g., low- or high-pass) filtering. One first simple approach to examine effects of this fact is to construct rhythmical patterns using several samples of each instrument. Experiences in a first pilot study with such "drum-machine-like" patterns suggest that individual note–to–note differences have a relevant influence on the overall character of the musical results.

1   Introduction

Whoever is experienced in listening to or playing traditional musical instruments will agree with the statement that their potential of emotional expression usually relies on complex variability of the sonic results in reaction to a musician's various input gestures. It has indeed often been noted that the sonic aesthetic fascination of an instrument is usually hardly captured in the quality of a single typical tone but also lies to strong degree in the dynamic behaviour of the abundance of various producible sounds. This observation is easily made in particular when dealing with electronic instruments based on playback of so-called "samples" of preexisting sound sources, as often used for the digital simulation of mechanical, electro-mechanical, or analogue-electronic instruments: sound production on the basis of sample-player simulations of mechanical instruments can easily lead to an overall result of the typical "MIDI studio sound" with its tendentially "inorganic", "static", or "cheap" character. The aspect of dynamic variation and reactiveness is therefore one of the hopes and goals connected to approaches of sound generation that are based on (physically inspired) models of processes of sound generation, and that are often denoted under the term of "physical" or "physically-based modelling". In sample-based sound generation, on the other

hand, strategies to introduce a dynamic quality to the instrumental behaviour
include the post-processing of employed signals by means of filters and amplitude
envelopes with dynamically varying parameters [1]. Mechanical sound-producing
processes, however, generally include much more complex dependencies between
characteristics of physical actions at the beginning of the causal chain (such as an
instrumentalist's gestures) and the final acoustic output. As an example, sounds
of an acoustic piano played at different strengths do not simply differ in a way
that may be fully described by amplitude scaling or low-pass filtering. Many
sample-based digital instruments therefore use "multisamples" of instruments
played at various levels of loudness. (For an example from the field dealt with
here, drum programming, see, e.g., [2].)

Inherent non-repeatability of sounds of mechanical instruments? While
variations of piano sound resulting from key strokes of different velocities or of the
sound of a drum struck at variable velocities have received some attention (and,
as just noted, have been considered by the use of multisamples), the question of a
possible perceptual relevance of unavoidable, involuntary indeterminacy in such
sounds has hardly been addressed so far. In more simple words: is it not possible
that indeterminable differences in, e.g., piano tones or drum sounds played with
identical intention on the side of the instrumentalist are an important factor for
the "overall sonic characteristic" of the instrument?

To establish the potential interest of this question it is worthwhile to note
that the mechanical systems of sound generation of traditional instruments in
their description as dynamical systems usually include strong non-linearities (see
e.g. [3]) so that even arbitrarily small variations in initial conditions or input
parameters may lead to strong and perceptible differences in sonic output. The
first author has been involved in the development of a sound synthesis algorithm
based on a simple model of impact interaction of solid objects. As an example, for
the observation just given, Figure 1 shows the vibration of a simple linear object
of two vibrational modes being struck periodically with equal initial velocity by
a point-mass (a sound event similar, e.g., to a mechanical door bell). Without
further explanation of the model by means of which this displayed waveform has
been generated (we here refer to dedicated publications [4]), it is noted that this
behaviour is very different — and *sounds different* — than a simple repetition
of identical copies of the signal created by one single impact. A closer look at
the displayed signal reveals that the different phases between two impacts also
differ more fundamentally than described simply by amplitude scaling or linear
(e.g. low- or high-pass) filtering. Given this complex behaviour of even a rather
simple dynamical system (derived from physical descriptions under involvement
of strong simplifications), we have to assume that for "real" mechanical systems
such as a drum or piano, the given observations are also, indeed even more,
relevant. A played drum roll must therefore have a very different auditory quality
than the periodic playback of the same drum sample (an example that has been
auditorilly tested by the authors and that will later be picked up again).
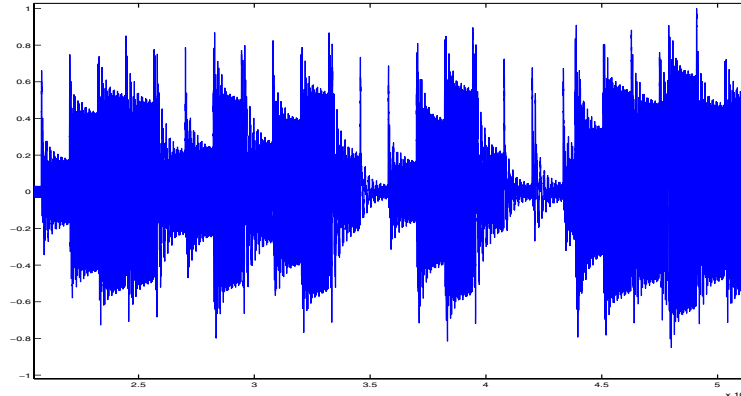
Fig. 1. Vibration (at one point) of a two-mode object struck periodically (at a frequency of approx. $20Hz$) by a point-mass

If we stick with the example of the drum, we must further consider that also many other initial conditions and parameters of the mechanical sound-emitting event beyond the initial state of movement of the membrane may be controlled by the player only up to a certain precision: examples are the velocity of the drum stick, the position at which the stick hits the membrane, the angle of the stick at the moment of contact (thus its "effective admittance"), and the force which the player exerts on the stick at the moment of contact. We must therefore take the possibility into account that even two drum strokes played at two freely chosen distinctive moments with the intention to sound the same, and with the drum being initially quiet, may actually sound clearly different. Figure 2 displays the signals of such two distinct snare drum strokes recorded under equal conditions and normalised for equal RMS power value. The two signals sound very similar, but could be distinguished auditorilly by expert listeners in a quick informal test. Their auditory difference is however not of the type achievable by amplitude scaling or linear filtering (a fact that is also reflected in the waveform displays).

From these considerations it seems convincing that the question given at the beginning of this paragraph is clearly worthwhile investigating in the case of drum instruments. Even for the piano, where variations of relevant parameters during sound generation (such as the ones listed above for the drum) are much more restricted by the mechanics of the instrument, it is imaginable that the differences between tones played with identical intention, and such that, e.g., the perceived loudness is equal, may be a perceptually (and musically) relevant factor. The work described in the following focuses on drum sounds for the reason of easier practicability and forms a first simple investigation towards the posed question.
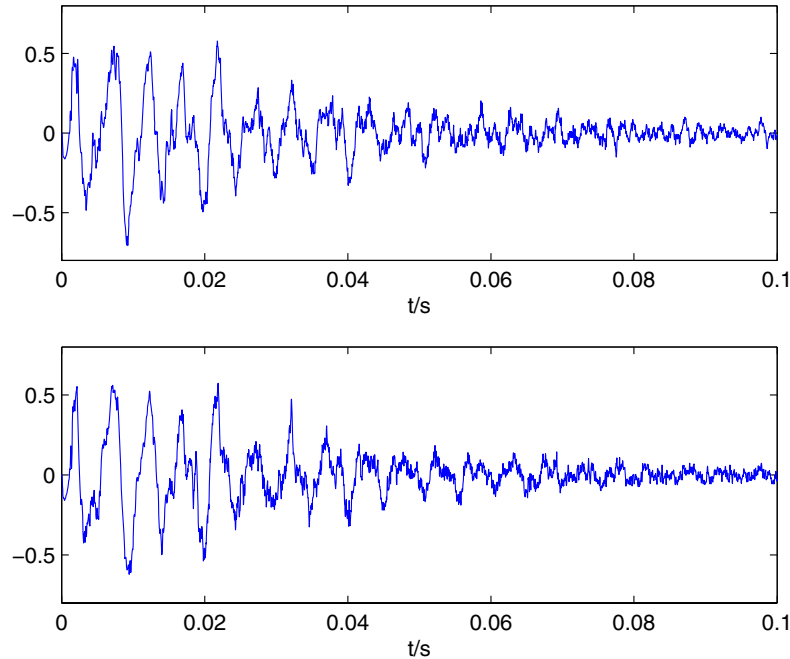
Fig. 2. Different snare drum strokes played with identical intention and normalised for identical RMS power value

2 A Simple Application of "Non-Repetitiveness" Considerations

In order to investigate the question described and motivated in the previous section we constructed drum-patterns from samples of bass drum, snare drum, and hi-hat strokes, in the way typically employed in a MIDI sequencer or drum computer. All patterns were then realised in two versions, one by using one sample of each instrument, just as in a conventional sample-based drum machine, and a second by using several samples of each drum instrument, distributed randomly among the programmed notes of the respective instrument. The different samples of each drum instrument were taken from previous recordings of the second author playing isolated notes on each instrument (bass drum, snare, and hi-hat). The single notes had been played with the same intention of loudness and playing technique to create a pool of potential drum samples for musical use. All tones of one instrument had been recorded with drum, microphone and recording room in the same constant configuration.

Individual differences beyond loudness variations? When listening to the different samples of one instrument from the recordings "in a row", or selectively comparing different such samples, differences are perceived (as already noted

above) but are subtle and very hard to describe verbally. Commonly understood attributes that might be used to describe some of the perceptual differences between single drum samples are small variations in loudness, and in rare cases in the length of the decay of the snare sound. In order to minimise an influence of these consciously notable aspects on the overall impression of the drum patterns — such as for example one single extreme snare stroke dominating the appearance of one version — two measures were taken: 1. Snare samples with noticeably short or long decay were not used in the pool of samples for random selection. 2. All samples of each instrument were normalised for equal RMS power level. With these precautions, the chosen samples of one instrument appeared of equal loudness to the authors.

In order to further focus on the relevance of individual sound variations that go beyond what may be captured by a loudness attribute, all notes in all generated patterns were scaled by a random factor distributed logarithmically uniform between $1/1.15$ ($\approx -1.2dB$) and 1.15 ($\approx 1.2dB$). The motivation for the choice of this exact factor of $+/-1.2dB$, which is close to the perceptive threshold for amplitude differences (see, e.g., [5]), will become clear in the following paragraph.

Drum rolls as "most delicate scenario". It has already been noted in the introduction (Section 1) that a drum roll must be expected to sound and look different than the periodic repetition of the sound of one singular drum stroke. The authors have constructed various versions of artificial drum rolls by triggering snare drum samples out of the pool of RMS-normalised samples at constant high rate, exactly 32th notes at a tempo of 120 BPM. The different versions consisted of

1. periodic repetition of one single sample,
2. periodic repetition of 2-18 different samples *in constant order*, and
3. constant–rate triggering of 2-18 different samples — *randomly*.

Figure 3 shows waveforms of three examples. Summing up the authors' informal auditory experimentation with the different created "roll" patterns, it can be noticed that

1. the periodically repeated single sample sounds very unnatural, almost not drum-like,
2. for several samples, periodic triggering in constant order results in audible periodicity in the pattern, which gets increasingly dominant the smaller the number of used samples, and is still disturbing even for 18 samples,
3. constant–rate triggering of samples in random order generally leads to auditory results closer to a played drum roll. As expected, results improve with the number of used samples. In the authors' informal experimentation, a number of four samples was found to form a minimum below which the overall sound becomes very unnatural.

Finally we examined the effect of additional amplitude randomisation with the general findings that
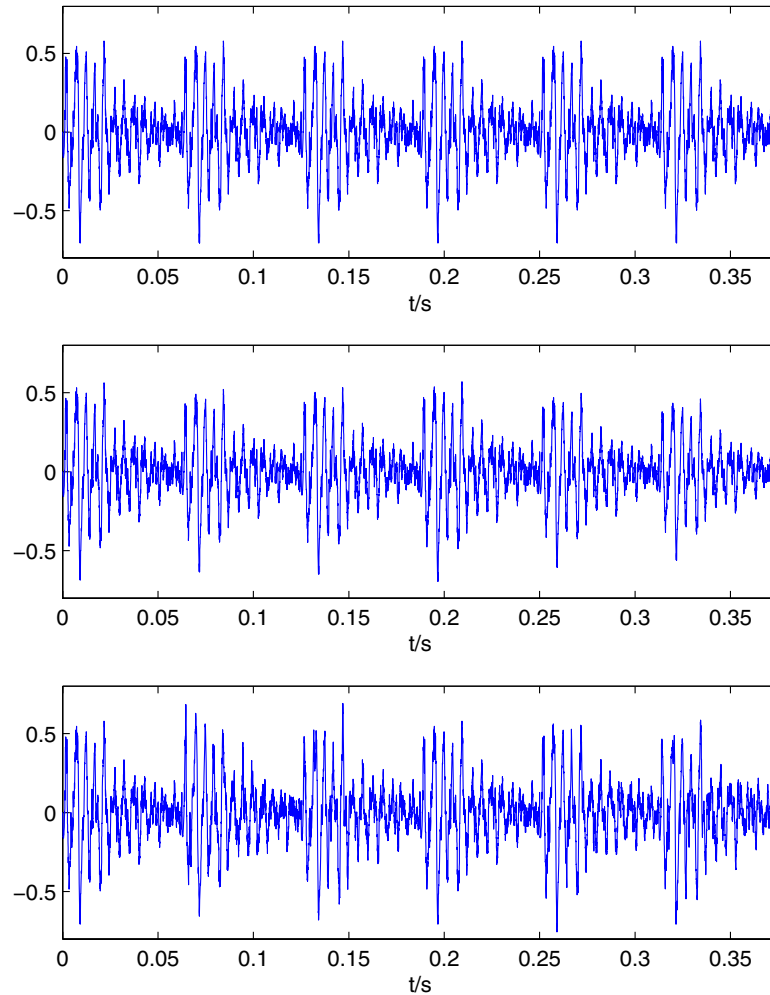
Fig. 3. "Drum rolls" created by repetition of one fixed sample (above), repetition of the same sample with random amplitude scaling of maximal $+/-1.2dB$ (middle), and constant-rate triggering of different samples of the same drum (below)

1. for periodic repetition of one sample, the sonic result gets somehow more natural up to a certain factor of randomisation above which the irregularity sounds rather awkward,
2. for the periodic repetition of several samples in fixed order, the timbral periodicities noted above remain also for "modest" amplitude randomisation, and
3. for randomly triggered samples slight amplitude randomisation has small auditory effect.

An amount of random modulation of individual amplitudes of samples between the maxima of 1.15 and 1/1.15 ($+/-1.2dB$) was chosen as leading to the most convincing sound results.

A "drum machine with sound variations". On the basis of the notions just described, a simple "sequencer" was realised that allows to program drum patterns of the three chosen instruments, bass drum, snare drum, and hi-hat in an arbitrary regular temporal grid, as with a typical "drum machine". Individual amplitudes of notes can be specified with 32-bit–float resolution and are randomised with an additional $+/-1.2dB$–factor as described above. Furthermore, each instrument (bass drum, snare drum, hi-hat) allows the specification of the number of samples in the pool out of which single notes are randomly taken. The random algorithm was enhanced such that no specific sample will appear twice in a row. The "drum machine" was realised in Matlab such that the temporal placement of drum sounds is accomplished with a precision of single time steps, i.e., precise up to approx. $0.023ms$ ($\approx 1/44100Hz$) at the used sample rate.

3   Pilot Study

In order to examine a potential auditory relevance of involuntary individual differences in drum strokes for musical character on a larger scale, a short pilot study was conducted with 12 test subjects listening to drum patterns generated by means of the experimental drum machine described in the previous section. Four drum patterns were programmed, somewhat of increasing complexity, each realised in two versions:

 – one "*static*", using only one sample per instrument and
 – one "*dynamic*", using eight samples per instrument.

The first pattern forms a rather simple, Disco-like beat at tempo 120 BPM, with a completely regular four–quarter note bass drum, and snare-strokes on beats 2 and 4. Pattern 2 is a slower, tempo 108 BPM, more complex "Funky Drummer"-like beat. Pattern 3 contains short 48th snare note rolls and is faster, tempo 130 BPM. Finally, still more complicated is pattern 4, a Drum'n Bass inspired beat with more drum rolls. All patterns are in 4/4 measure with regular 16ths as smallest rhythmical subunit (with the exception of occasional 48th note "rolls"). Figure 4 shows the first and fourth example in schematic notation.

3.1   Experiment

At the beginning of the test, each subject was asked to listen to the four drum patterns, each in the two versions. The order in which the two versions were presented was fixed for each subject but swapped in-between subjects. Each version was preceded by 4 seconds of pink noise (followed by 1 second of silence) in order to "erase" any immediate memory of preceding patterns. As an example, stimuli were presented to the first subject in the order: 1. pink noise, 2. beat 1,
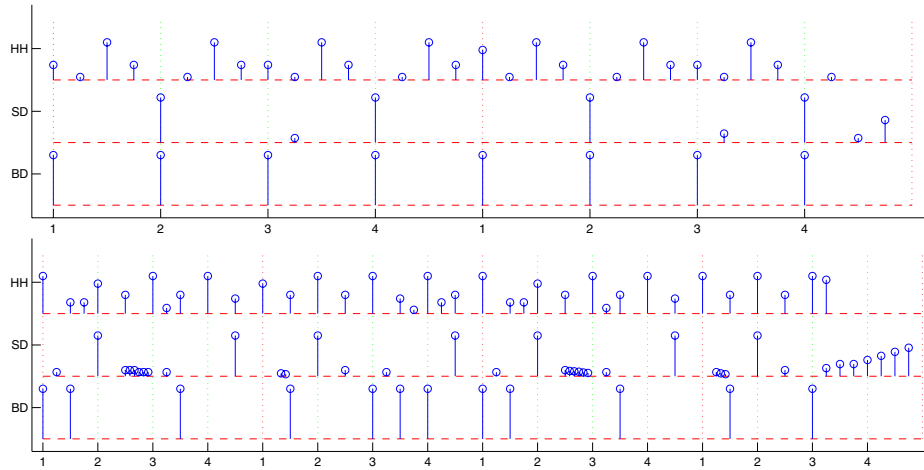
Fig. 4. The first (top) and last (bottom) of the drum patterns used in the listening experiment. The horizontal unit is given by quarter note beats.

dynamic version, 3. pink noise, 4. beat 1, *static version*, 5. pink noise, 6. beat 2, *dynamic version*, 7. pink noise, 8. beat 2... Before listening, subjects were told that they would hear four drum patterns, each in two different versions, separated by noise.

After having listened to the series of drum patterns, subjects were asked five questions, the first two of which aimed at subjects' immediate spontaneous impression: "In what respect do you feel does the first version of each beat differ from the second one?", "Which version of each beat, the first or second, did you generally like more?".

Question 3, as question 2, again aimed at the personal preference of which version the subject thought sounded better, but now the subject could (if desired) listen again to the pair of each beat before answering. The same procedure applied to the questions 4 and 5: "Which version of the beat sounds 'more natural'?" and "Which version do you feel sounds more like a human drummer rather than a machine?". All questions were read to subjects and explained in case of difficulties. In particular, subjects were told that they did not have to answer to the questions. Twelve subjects participated in the test, three female and nine male, aged between 25 and 35, with various musical experiences and interests.

3.2   Results

Spontaneously perceived difference. As a first general observation, 5 of the 12 subjects, after having listened once to the series of pairs, found it hard to spontaneously state or at least to verbalise any perceived difference. Exact answers of these 5 subjects included "no difference noticed" (one subject) and

doubtful guesses without obvious relation to the real phenomenon such as "more reverb?", "more dull?, faster?" or "further away?". Of the remaining 7 subjects, one described the *dynamic version* as being "more dynamic, as if the drummer was more involved in dosing the strength of his strokes to make the beat more energetic", one labelled this version as "more natural", the other as "more artificial" and one stated the *static version* being "unnatural". One subject described the differences in the pairs as "the dynamic fluctuations in the snare drum being differently strong" without being able to give a general tendency from the *static* to the *dynamic version*. The three remaining subjects gave more generic attributes, concretely: the *static version* "sounding cleaner", "more pushing", the *dynamic version* "sounding stronger". No subject, including those with some musical and listening experience, was able to actually find out the mechanism behind the generation of the pairs of beats (namely one version being built from different samples and one containing only one sample of each instrument).

Personal aesthetic preference and "naturalness" judgements. Asked for a spontaneous preference after first listening, 5 subjects opted for the *dynamic version*, 3 for the *static* one. The remaining 4 subjects had no preference. When allowed to listen again to single pairs of beats under the aspect of preference, in few cases some subjects rendered their responses more precisely. Without going into details, the general tendency here was towards the *dynamic version* being preferred. One subject that had initially preferred the *static version* now preferred the *dynamic* one for all four beats, as well as another subject that had initially not stated any preference.

Answers to question 4 correlated strongly with those to question 3, i.e., in almost all cases subjects now labelled the version as more "natural" that was previously liked more, whereby the *dynamic version* now clearly dominated. The same overall tendency was finally seen for question 5 where 5 of the 12 subjects labelled the *dynamic version* as "sounding less machine-like" than the *static* one and only one subject said the opposite for the first pattern.

While this test and the presented results are of course preliminary, it can however generally be stated that a difference between the *static* and *dynamic* version is indeed perceived, although subjects often have difficulties describing their impression. There is at least some evidence that the *dynamic* version is preferred in average. Of course, these remarks must be considered as a first exploration and will have to be pursued deeper in extended experiments.

4   Conclusions

The presented work forms a first exploration on the question of a potential relevance of inherent variations in individual sounds of mechanical instruments. Despite the very simple approach used here to take into account possible variations in drum sound for the creation of stimuli patterns and the very simple design of the pilot study, some effects are seen. Strong hints have been found that unavoidable differences between drum sounds of essentially equal perceived

volume and played with identical intent may be clearly relevant for the perceived character of a musical phrase. From these observations it appears well possible that the conventional practice of constructing rhythmic patterns from single samples of percussive instruments may be one important factor responsible for what is commonly perceived as the typical "MIDI studio sound". A closer examination of the phenomenon may form a basis for new tools for musicians with extended expressive possibilities.

References

1. Massie, D.C.: Wavetable sampling synthesis. In: Kahrs, M., Brandenburg, K. (eds.) Applications of Digital Signal Processing to Audio and Acoustics. Springer, Netherlands (2002)
2. `http://www.fxpansion.com/`
3. Smith III, J.O.: Physical modeling synthesis update. Computer Music Journal 20(2), 44–56 (1996)
4. Avanzini, F., Rath, M., Rocchesso, D.: Physically-based audio rendering of contact. In: ICME 2002 Proceedings. IEEE International Conference on Multimedia and Expo., vol. 2, pp. 445–448 (2002)
5. Zwicker, E., Fastl, H.: Psychoacoustics. Facts and Models, 2nd edn., Berlin (1999)