GranularSpectrals – Spectral manipulation for Advanced Granular Synthesis

Fabian Töpfer
Lars Engeln
Rainer Groh
fabian.toepfer@tu-dresden.de
lars.engeln@tu-dresden.de
rainer.groh@tu-dresden.de
Technische Universität Dresden, Institute of Software and Multimedia Technology
Dresden, Germany

ABSTRACT

Spectral manipulation has increasingly established as a common method in audio editing. Based on the analytical visualization of audio material as a spectrogram, it offers interactive methods that are visually comprehensible. Consequently, it can be used to extend methods of music production by discovering novel techniques or by revisiting known ones. Thereby, granular synthesis has high potential in gaining functionality when thought and used in a spectral domain interface. Here, granular spectral synthesis is proposed and how it takes advantage is discussed.

KEYWORDS

granular synthesis, spectrogram, visual audio design

1 INTRODUCTION

Already Gabor [7] described the later well known spectrogram. Furthermore, Gabor theorised the decomposition of audio signals into *acoustic quanta*, which are referred as *Gabor atoms* (compare [**?**]). They are the elementary signals which occupy the smallest possible area within the spectrogram [7]. The discussion about the smallest building blocks regarding to time and frequency of sound was the earliest begin of granular synthesis. Thereby, the foundation of granular synthesis is the observation of micro-structures in audio signals [5]. These micro-structures are also called *grains* for the composition of sound objects [**?**]. Extracted grains building up a so-called *grain pool*, from which grains are then combined into new structures [1]. They can be arranged in a new, overlapping way. An envelope is applied to the of the grains, which ensures a homogeneous superimposition of the grains [5]. The new arrangement of the grains is specified by the scheduler of the synthesis. There are three different scheduling strategies (synchronous, quasi-synchronous, asynchronous), which are characterized by their temporal distribution of grains [1]. While synchronous may induced a periodic pattern by accidentally canceling out with itself and asynchronous may lead to gaps in the signal, quasi-synchronous can guarantee a consistent signal. Numerous systems and different interfaces for granular synthesis have emerged [2, 8–10, 14, 15].

Normally, granular synthesis is computed in time-domain, but also a frequency-domain approach has been proposed (see [6]), but no new interactions with-in the frequency-domain have been discussed. This is why this paper explores with the transition of granular synthesis into the frequency-domain novel interactions and expressions of advanced spectral granular synthesis.

2 DESIGN SPACE

The spectrogram offers more possibilities for arrangement and manipulations [3]. Whereas conventional time-domain granular synthesis rearrange the grains on the time axis, spectral granular synthesis can make use of the second dimension and in addition rearrange the grains on the frequency axis as well. Grain arrangements can vary from perfectly aligned to randomly distributed on the frequency axis, resulting in a completely new auditory character. Additional transformation are also possible. The spectral data of the grains can be manipulated in additional ways, like rotation and scale (compare [3]). A representation of the granular synthesis' parameters thought in spectral domain are shown in Fig. 1 . The parameters, which determine a grain, are its size and spray. The size corresponds to the length of the audio material the grain has and can be independent for each grain and/or grain pool. The spray is the area with-in the source from which the grain is extracted. If the spray equals the size, always the same material will be extracted. The processing in spectral domain also allows to manipulate the frequency as well as the time. So, having a size and spray on the frequency axis inducing a randomized low-/high-pass is possible, but here we keep the full spectrum. With a small spray, i.e. using only grains from a relatively narrow area of the source, the stream has a very homogeneous sound. Whereas with a larger spray, the sound becomes more varied and colourful, while the impression of a sound collage is created. Additionally, if the spray is small, but the corresponding area of the source shifts over time more effects can be achieved. So, with a small enough spray and the corresponding source area moving in just the correct speed, granular synthesis creates a time stretch effect.

The composition of grains is influenced by the density, interonset and duration. The density is the number of grains placed in a specified time duration. The duration of the composition represents the total length of the composition of grains being generated. The

MuC'22, 04.-07. September 2022, Darmstadt

https://doi.org/10.18420/muc2022-mci-ws03-397

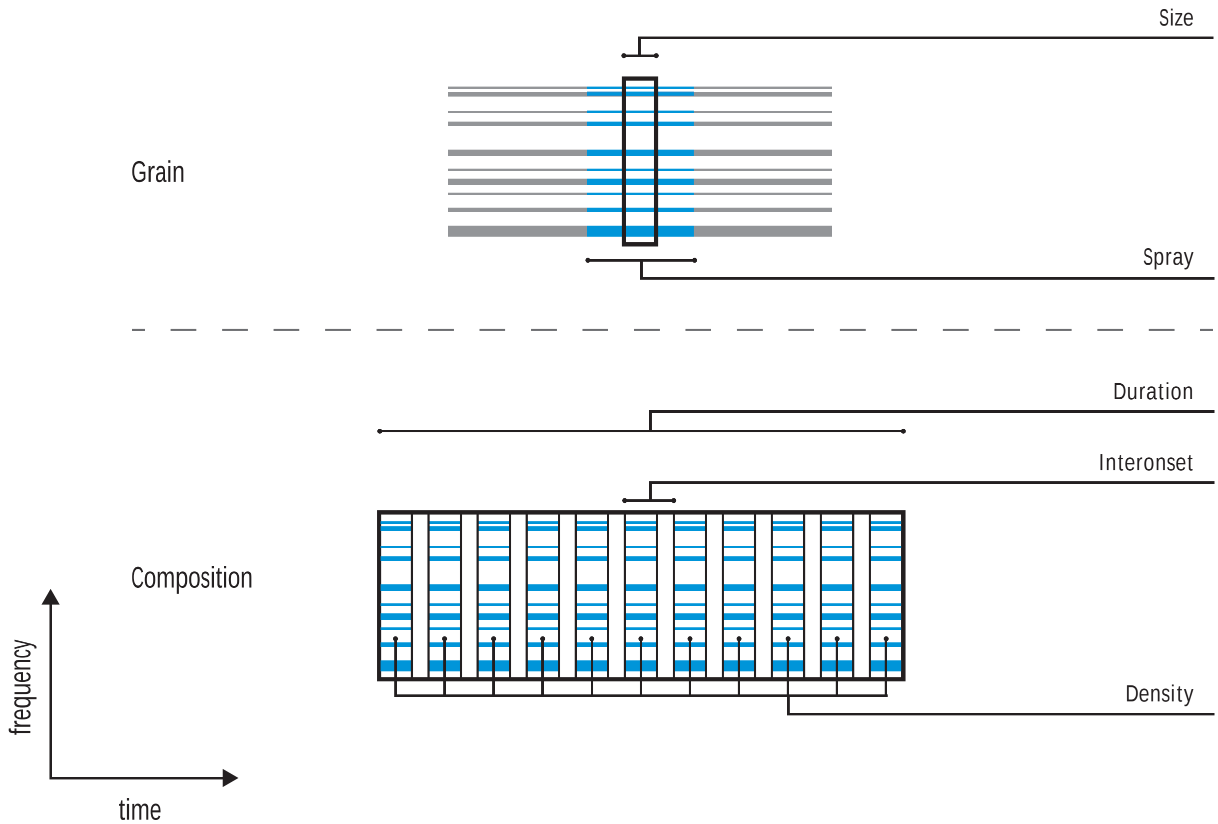

Figure 1: Parameters of granular synthesis in spectral domain. For visual clarity, grains are not shown being overlapped, as it would be the case with interonsets typically shorter than the grain size.

interonset is the time interval between successive grains. The whole composition can be changed in time and frequency position, as well as each grain can be moved individually.

These parameters lead to an hierarchy, since the granular composition of one sound consists of several grains and each grain consists of audio material. One sound, so one composition of grains is dealt as group of grains. The parameters for composition influence the grains on a higher and structural level by influencing the relations between the grains. Speaking for one sound, the parameters of the composition can also be considered global and the parameters of a grain local. Considering a potential large set of grains, a selection of local parameters of the grains should be adjustable both, global and local. Multiple granular compositions of different sounds can be edited alongside each other, creating a bigger overall composition in time and frequency.

The transformation of the signals into the frequency-domain affects the realisation of the granular synthesis itself and results in special features. Grains are spectrum slices of the spectrograms, which are combined to build up a composition. For the FFT a window function is applied to the discrete values of the signal, which acts like an envelope weighting the signal locally. So, the necessity of windowing and overlapping as in time-based granular synthesis comes by nature of the FFT. Especially, as "sharp temporal edges" in the spectrogram (high spectral flux) are smoothed out with the windowed reconstruction, the signal will always be continuously differentiable. Although, short sharp events (smaller than the window) will also be extremely smoothed and diminished. But thereby, artefacts that occur in time-domain granular synthesizes with short grain sizes at approximately 10ms to 15ms (cf. [1]) do not occur here.

By looking at the scheduling strategies, some side-effects sometimes considered as disadvantages can be used creatively. Synchronous scheduling may result in unwanted cancellation of the produces stream. Thereby, quasi-synchronous produces streams with a very high grain superimposition, which can result in a very continuous and even sound [12]. A reduction of the grain size inevitably causes the generation of noise [12]. Pure asynchronous scheduling leads to holes in the stream and on the other hand, to *clouds* of overlapping grains at non-regular intervals. By exaggerating them, different sound textures can be achieved in the manner of granular synthesis. For *stream* and *clouds* they can overlap strongly in real use). While the *stream* texture is the common smooth outcome, the *stutter* texture uses a large distribution of the grains with higher interonset than grain size to create a sound that has gaps at regular intervals. The *cloud* texture uses a highly asynchronous scheduling, thus pseudo-random distribution of the grains in time, resulting in swelling and shrinking as well as silent regions.

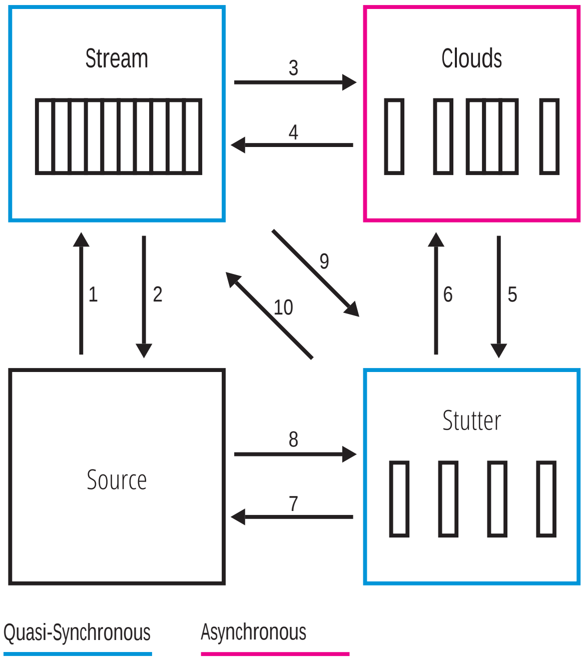

Figure 2: Relations of sound textures. (1) Increase of superimposition, (2) Decrease of superimposition, (3) Variation of density, (4) Leveling and increase of density, (5) Leveling and decrease of density, (6) Variation and increase of density (7) Increase of superimposition, (8) Decrease of superimposition, (9) Decrease of superimposition, (10) Increase of superimposition.

3 ADVANCED SPECTRAL GRANULAR SYNTHESIS

A granular synthesis with advanced spectral interaction was created and implemented. *GranularSpectrals* is available as a module in the VisualAudioDesign (VAD) [3] environment and is thereby a C++ application based on libCinder[1] and OpenCV.

The individual steps during the synthesizing process as well as a selection of granular features are divided into tools, which the user can activate or deactivate. Each tool has a specific interaction, which are further examined. Is the selection tool activated, the user selects a region in the spectral domain to create a source entity (see Fig. 5). Multiple sources can be created this way for further processing and granular synthesis. In a following step, the user can choose between a quasi-synchronous and asynchronous granular synthesis. The selected source entity is used for placing the newly generated composition entity. This placement is done via a drawing tool, which enables the user to freely draw a line in the spectral domain. The line serves as a reference to the fundamental tone or the lower end (0Hz) of the grained spectrograms of the individual grains with-in the composition as guidance of grain placement. As represented in Fig. 3, grains can follow the line by each spectrum column which manipulates the pitch with-in a grain, or a grain aligns with the middle mass to the line. After drawing is completed, the composition is automatically generated. This way of creation is identically for both scheduling strategies.

With the time-stretch tool, the user can select an area, similar to the selection mode. This area can be scaled horizontally, see Fig. 5. The spectrogram become compressed or stretched. In merge mode, the user can connect two compositions and/or sources. This is achieved by drawing a line from one composition to another. A transition is then generated between the two compositions (see Fig. 5) creating the opportunity for creative transitions (cf. [4]).

Three levels of manipulations are proposed, which are used to change various parameters of granular synthesis. This hierarchy is not introduced via a tool, but forms a paradigm of interaction that forms the frame of the workflow. The aim of this hierarchy is to satisfy the needs of diverse users and to enable clear and finely granular manipulation. With that, it follows Shneiderman's mantra: "Overview First, Zoom and Filter, then Details on Demand" [13].

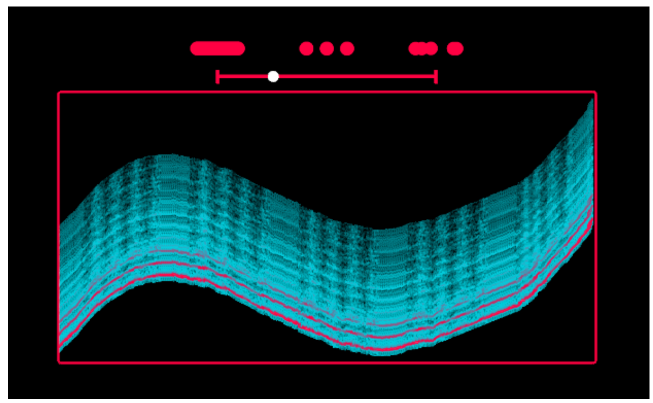

Figure 3: Example manipulation in level 1 of the sounds texture by influencing the scheduling on one slider.

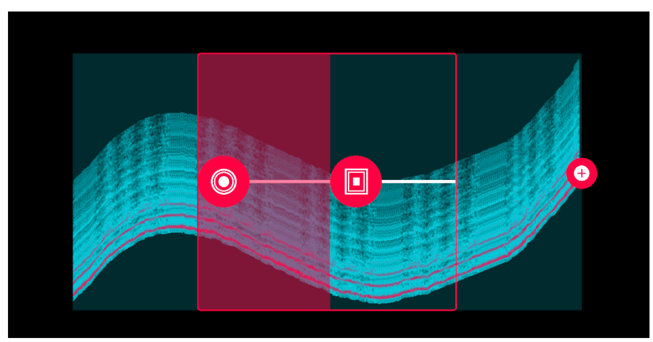

Figure 4: Example manipulation in level 2 – changing the grain size to a high value.

Level 1 regards the type and strength for creating different textures – so the type of scheduling, and the interonset offset and strength of pseudo-random position. It offers the possibility to change these parameters with a single slider input (see Fig. 3). The goal is to switch smoothly between the different texture (see Fig. 2). At level 2, all global parameters of the grains from one composition

[1] https://libcinder.org/

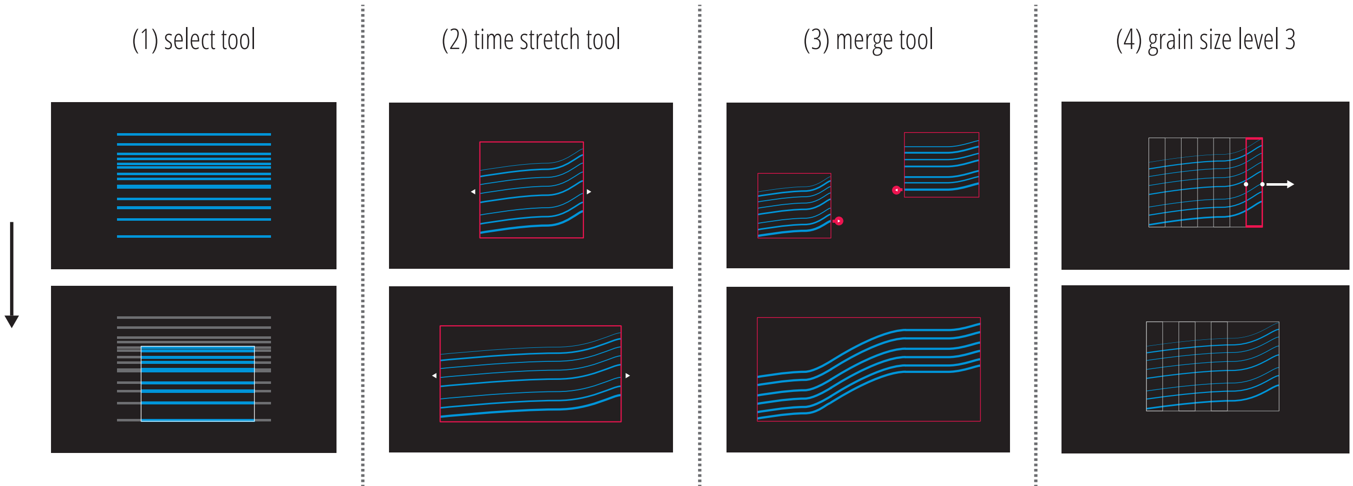

Figure 5: (1) Selecting a source entity while the selection tool is activated, (2) Using the time stretch tool, (3) Using the time merge tool, (4) Example manipulation in level 3 - Adjusting the size of an individual grain.

can be manipulated separately via specialized inputs (see Fig. 4). Thus, a more precise manipulation of the grain size, interonset, and with that the density is possible. Level 3 in the manipulation hierarchy is characterised by the local manipulation of a grains. The aim of this level is a very fine granular manipulation of the composition. The user can adjust the parameters of a single grain. These manipulations can be applied to a selection of grains that contains at least one up to all grains of the composition. The remaining grains of the composition remain untouched, as seen in Fig. 5. This stage is only useful if the number of grains within the composition is manageable or a suitable view is chosen that guarantees a clear representation even with a large number of grains. Also, if the general structure on a higher level is changed, individual changed grains might disappear in the current stage of the implementation. If unique grains are changed, these should not be influenced by the scheduler.

4 DISCUSSION

The implemented prototype is a proof-of-concept for novel interaction with granular synthesis in the spectrogram. Manipulation of a single grain might not as useful for every editing task, as there are the not examined ways how the scheduler should behave in this case. If a single grain should be extensively unique in position/rotation/scale, then it would be better, to introduce this grain as new sound besides and separated from the granular composition itself. By doing that, the scheduler has the freedom to complete regenerate the sound, but the special grain as additional sound is reserved. The manipulation of a single grain is only really suitable, if some grains from the semi-automated scheduling can be optimised for a better sounding collage. This is the case, if some transformations, especially extreme rotation and scale with-in the thresholds, are standing out auditorily. The general workflow should be top-down anyway. If the general shape and texture is sufficient, then grains shall be tweaked. If the general shape is changed due level 1, the whole sound is recreated and prior changed grains might fit in better again. Therefore, all locally changes to grains should be vanished if the scheduling changes noticeably.

Way more interesting would be a study on the creative use of this advanced spectral granular synthesis, to weight which tools and texture are more likely to be used (in comparison to 2). With that information, the workflow and the granularity of interacting with the scheduler can be tailored to this application. In addition, the investigation of different styles of visual collages might yield novel way for auditory collages. Also a qualitative approach for interacting should be considered. Thinking of smooth streams up to collages which are rich on contrast.

5 CONCLUSIONS

Advanced possibilities for granular synthesis in the spectrogram are explored and possible interactions are proposed. Various visual-audio interactions, as well as special features for the application of spectral granular synthesis were considered. For better ease-of-use, the interface was split into different level with different granularity of manipulation and orchestration of the grains. The next steps include optimisation and extending the prototypical implementation. As well as evaluating the interface in a creative use-case. Based on this work, different sound synthesis techniques could be examined in the spectral domain, too. Transferring this system in the three-dimensional space and using mixed reality could also be beneficial for shaping sound and will need special investigation.

REFERENCES

[1] Ross Bencina. 2001. Implementing Real-Time Granular Synthesis. August (2001), 1–17.
[2] Chris Carlson and Ge Wang. 2012. Borderlands : An Audiovisual Interface for Granular Synthesis. *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression* (2012), 260–263.
[3] Lars Engeln and Rainer Groh. 2019. VISUALAUDIO-DESIGN – TOWARDS A GRAPHICAL SOUNDDESIGN. (2019), 1–8.
[4] Lars Engeln, Dietrich Kammer, Leon Brandt, and Rainer Groh. 2018. Multi-Touch Enhanced Visual Audio-Morphing. In *Proceedings of the International*

Conference on New Interfaces for Musical Expression. Zenodo, 152–155. https://doi.org/10.5281/zenodo.1302711

[5] Andy Farnell. 2008. *Designing Sound, Excerpt*.

[6] Stefano Fasciani. 2018. Spectral granular synthesis. *ICMC 2018 - Proceedings of the 2018 International Computer Music Conference* July (2018), 99–103.

[7] Dennis Gabor. 1946. Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering* 93, 26 (1946), 429–441.

[8] Carl Gutwin, Oliver Schneider, Robert Xiao, and Stephen Brewster. 2011. Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 85–94.

[9] Sile O'Modhrain and Georg Essl. 2004. Pebblebox and crumblebag: Tactile interfaces for granular synthesis. In *Proceedings of the 2004 conference on New interfaces for musical expression*. Citeseer, 74–79.

[10] Leonard J Paul. 2017. *An introduction to granular synthesis in video games*. Routledge.

[11]]Roads Curtis Roads. [n. d.]. *Microsounds*.

[12] Curtis Roads. 2012. From Grains to Forms. *Proceedings of the International Symposium Xenakis* (2012), 1–40.

[13] Ben Shneiderman. 1996. The Eyes Have It : A Task by Data Type Taxonomy for Information Visualizations. (1996), 336–343.

[14] Barry Truax. 2003. Music and science meet at the micro level: Time-frequency methods and granular synthesis. In *Music Viva conference, Coimbra, Portugal*.

[15] Anil Çamci. 2018. Graintrain: A Hand-Drawn Multi-Touch Interface For Granular Synthesis. (Jun 2018). https://doi.org/10.5281/zenodo.1302528