GRANULAR SYNTHESIZER PROTOTYPE: GRAINSPAT

Frederico Pereira 440600239 (SID)

Final Written Review, DESC9115, 2016
Graduate Program in Audio and Acoustics
Faculty of Architecture, Design and Planning, University of Sydney

1. INTRODUCTION

Granulation is a particular application of the Granular synthesis technique, where a given recorded signal in digital format is broken apart in small elements, which can then be manipulated and reassembled to form an output signal. The technique has first been recognized by Xenakis in 1971 as a powerful musical application, and has since been used in countless pieces, by pioneering composers like Truax, Brummer, and, in more recently produced pieces by Fennesz, Kieran Hebden or Daniel Lopatin.
A presented prototype is described in this document. The prototype applies the granulation algorithm for a given input audio signal, and outputs up to four channels of audio. The signals in the different output channels can be decorrelated, in order to achieve a superior subjective spatial impression.
This prototype is presented in software form, as a group of Matlab functions and script.

2. GRAINSPAT IMPLEMENTATION

2.1. Basic features

The GrainSpat algorithm breaks an input signal into a number of elements ("Grains") of user defined duration, that can range between an interval of minimum and maximum, then, these elements can be assembled in a chosen sequence to form an output signal according to a user defined density (in Hz).

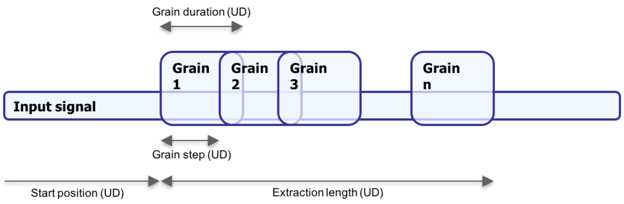


Figure 1. *Granulation of input signal. UD denotes User Definable parameter.*

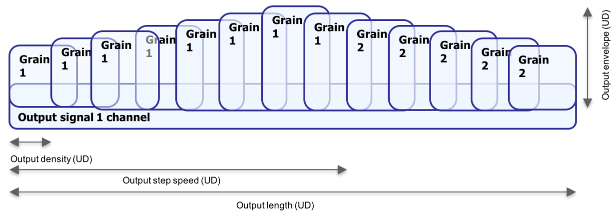


Figure 2. *Composition of an output channel. UD denotes User Definable parameter.(A forward arrangement with step speed =70% is represented)*

The grain routine is mathematically described by:

$$y(n) = \sum_{k} a_k g_k(n - n_k)$$

2.2. Parameters description

2.2.1. *Input signal and channelsIN*

The input signal is a digital audio file chosen and loaded into the routine using a standard Matlab function. The number of channels of the signal has to be defined for processing, generally the input signal will be a one channel (mono) or two channel (stereo) file. Definition of channels is done through "channelsIN" parameter.

2.2.2. *Startposition and Extraction length*

These parameters define the region of input to be used for granulation:

- Starposition, as percentage of total input signal duration indicates the first sample for extraction
- Extraction length as percentage of total input signal indicates the duration of signal, from startposition, used for extraction.

Note that the values indicated by startposition plus extraction length have to be less than 100% (less than total input signal)

2.2.3. *Read Type*

Read Type parameter defines the order of extraction from input, this is as well the order in which the grains will be placed in the output. It can be set to the following:

- Fixed (=0), the extracted grain is always the same.

- Forward sequence (=1), the extraction progresses forward in input time from defined start position
- Backwards sequence (=2), the extraction progresses from the end of extraction length towards startposition.
- Random (=3), extracted grains will be picked randomly from startpositon through extraction length.

2.2.4. Step

This parameter is set in number of samples. It regulates the number of samples that separates one grain to the following in the extraction length. Note that if the step is smaller than the grain duration, an overlap of grain signal results.(as in the case represented in Figure 1)

2.2.5. Stepspeed

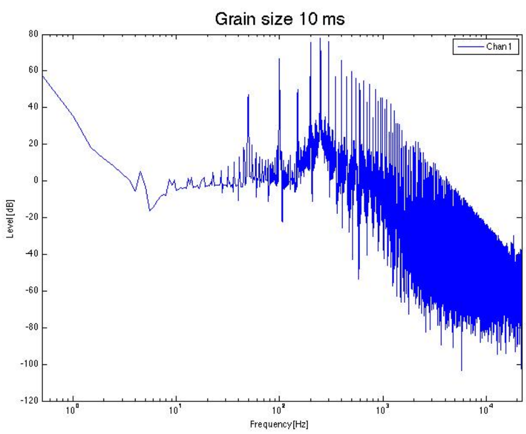
Stepspeed is set in percentage values, and regulates the number of same extracted grain repetitions placed in the output until the next extracted grain is placed. This parameter works in conjuction the read type parameter, and it's not applicable for the fixed read type. The percentage value represents a maximum (100%) and a minimum (0%), these maximum and minimum stepspeed values are predetermined in the routine (i.e not user definable) and limit a range of useful values.

2.2.6. Lout

Defines the duration, in miliseconds, of the output signal.

2.2.7. Grsize

Defines the duration of each grain in miliseconds. This parameter has major influence in the character of the output. The grain has information of the input signal in the time domain as well as in frequency domain. The shorter the duration of the grain, the widest the bandwidth of the output will be. Grains with durations shorter than input signal period still represent the pitch of the input, by with modulation products and noise. (C. Roads, Microsound).



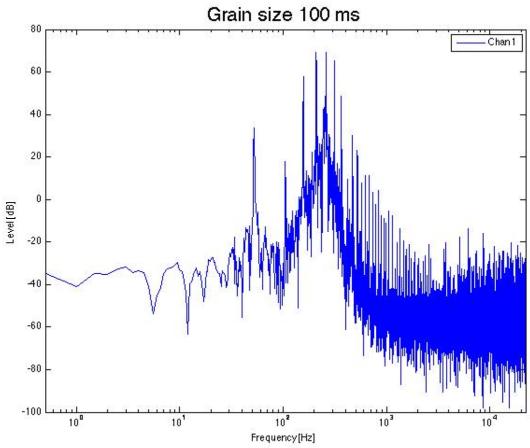
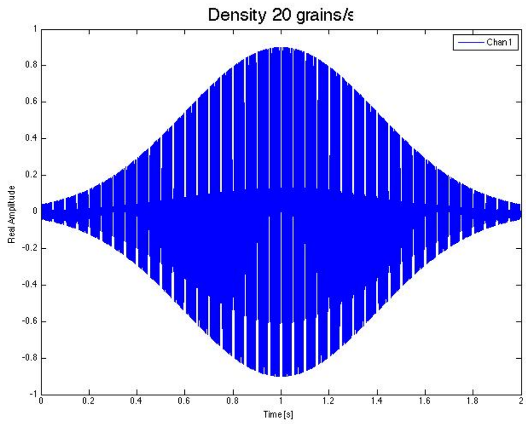


Figure 3. *(a) Spectrum of output with Grain duration of 10 ms from an input sine 800 Hz sine wave; (b) Spectrum of output with Grain duration of 100 ms) from an input sine 800 Hz sine wave. Wider spectrum using Grain with 10 ms is visible.*

2.2.8. Density

With values representing number of grains/second in the output, this parameter, has also a major impact on the output character. With low values resulting in a blocked, sparsed sound, and higher values in a dense and texturized sound.
It is encouraged to play with the combination of grain duration and density to achieve different character outputs.



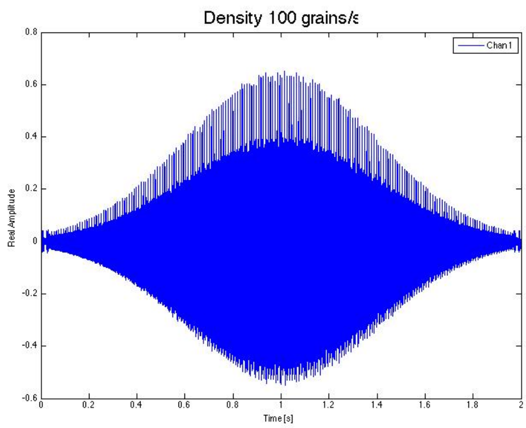


Figure 4. *(a) Time domain representation of output with density of 20 grains/s from an input sine 800 Hz sine wave; (b) Time domain representation of output with density of 100 grains/s from an input sine 800 Hz sine wave. Sparse sound of the 20 grain/s is visible.*

2.2.9. Window (W)

Sets the amplitude envelope shape for each extracted grain. This parameter controls the fade-in and fade-out curves, with values between 0 and 1 resulting in cosine curves of a Tukey window (flat region window), a value < 0 yields a rectangular window, and a value >1 a Hann window (no flat region).
The envelope shape has an effect on the spectral content of the output signal, with rectangular envelopes introducing clearly noticeable artifacts. The envelope can be regarded as an amplitude modulation of the output, creating additional sidebands on its spectrum (additional frequencies, above and below the grain frequency).

2.2.10. Spatial effects parameters: channels output

The number of channels to be output is defines by '*choutput*' parameter, where the user can chose up to four channels (for the current prototype version).

2.2.11. Spatial effects parameters: grain size variation

For each channel, the grain size can be randomly changed in between an interval defined by "*Gdmax*" and "*Gdmin*". The larger the interval, the larger the variation between signals.

2.2.12. Spatial effects parameters: density variation

This parameter accepts percentage values. For each channel, the density of grains in the output can be changed between no variation (0%), to larger variations (100%). A value of 100% varies the preset density in 100% (it can be max of double or min, ofone grain/second). Small values yield a subtle variation, fitting the purpose of spatialisation of the output. With larger values the signal starts to be easily discriminated between each channel output.

2.2.13. Effects

The output signal can be low pass filtered, this is done using a Matlab built in function that processes the signal through a FIR filter with 32 taps. The user defines the cut off frequency.
An artificial reverberation algorithm is also available to process the signal, this algorithm consists of four feedback comb filters and three all pass filters. The user specifies the maximum reverb time for the comb filters. The algorithm processes up to five reflections.

2.3. Routines mathematical description

2.3.1. Synthesis routine

The synthesis loop creates the output from extracted grains, it can be basically described by the following expression:

$$y(n) = \sum_{k}^{i} a_k g_k(n - n_k)$$

Each time the loop is called, a line array "Y" (output signal, initially) is populated with the grains retrieved from "x" (the input signal). On each iteration, the grain is summed to the "Y" array.
This is the algorithm's main routine. It processes the input signal and synthesizes the output according to the defined parameters.

2.3.2. Extraction routine

This routine applies the grain retrieval from the input signal array "x". The matlab function "tukeywin" is used to define the envelope window shape .
The extraction routine can be described by the following expression:

$$g_k(i) = x(i + i_k).w_k(i)$$

3. EVALUATION

The implemented granulation offers a number of parameters than can be user controlled. These allow for different types of processing, resulting in different sounding outputs. Of special consideration is the ability to spatialize the output sound by defining a change in parameters for each channel output. For the following examples and descriptions, a two channel output will be used.

3.1. Granulation spatial performance with density variation

As an example audio signal of 2 channels output, with no spatial processing applied, the file "dnsty_spatvar_0.wav" can be auditioned.
Spatial processing has been applied to the same input, definig a variation of density parameter for 1%, 2%, 5% and 20%, the respective audio files are: "dnsty_spatvar_2.wav", "dnsty_spatvar_5.wav", "dnsty_spatvar_10.wav" and "dnsty_spatvar_20.wav".

Subjective perceived spatial impression is generally well described by two attributes, Listener Envelopment (LEV) and Inter-Aural Cross Correlation (IACC). IACC is suggested as a predictor for Apparent Source Width (ASW). (Manson and Ramson, 2002). With a lower value of IACC translating to a predicted greater ASW.

Figures below show a comparison of IACC for the different variation degrees of density parameter.

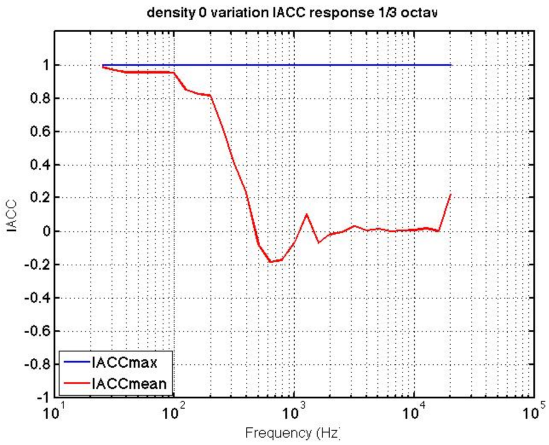


Figure 5. *IACC for no variation in density.*

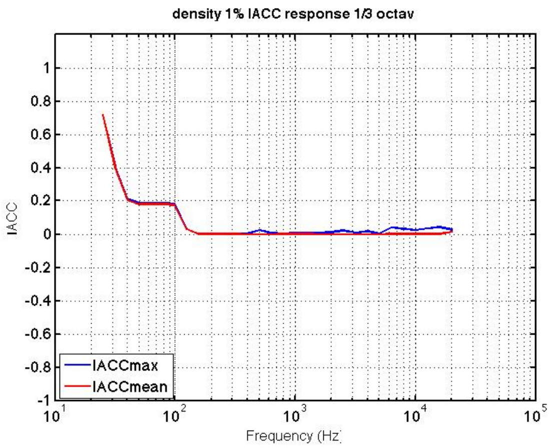


Figure 6. *IACC for 1% variation in density.*

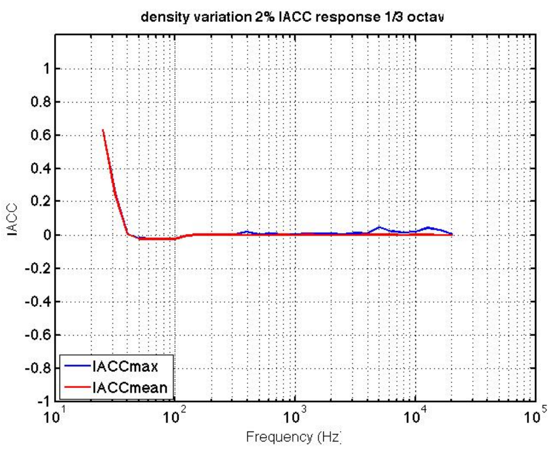


Figure 7. *IACC for 5% variation in density.*

As seen from the graphs above, the variation in density of 1% and 2% result in mean values of IACC very close to 0.
From auditioning the audio files, it can be inferred that values of density variation larger than 5% result in sounds being discriminated in the two channels, no longer contributing to spatial impression.

3.2. Granulation spatial performance with grain size variation

Relying on variation of grain size for decirrelation, the following examples have been processed:

- Variation between 30 and 70 ms (*"Gsize_spatvar_30_70.wav")*
- Variation between 20 and 80 ms (*"Gsize_spatvar_20_80.wav")*
- Variation between 10 and 90 ms (*"Gsize_spatvar_10_90.wav")*
- Variation between 10 and 200 ms (*"Gsize_spatvar_10_200.wav")*

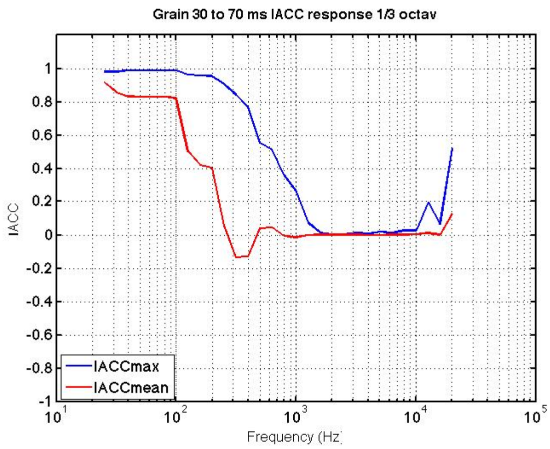


Figure 8. *IACC for 30 to 70 ms grain size variation in density*

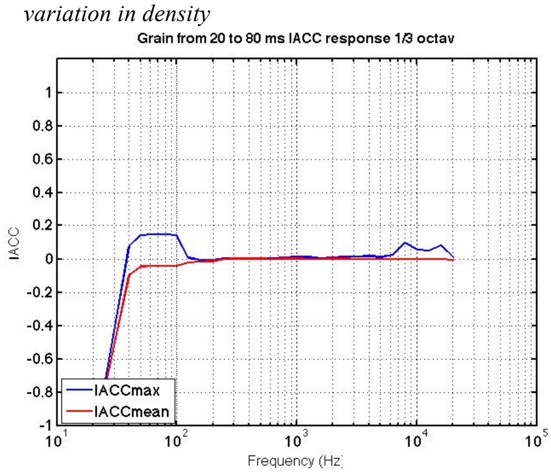


Figure 9. *IACC for 20 to 80 ms grain size variation in density*

3.3. Granulation spatial performance with random grain extraction.

With random extraction of grains from the same input, the different channels will logically have a different composition, these can lead to more or less successful spatial impression.

3.4. Creative use evaluation.

In order to evaluate the applicability of the proposed algorithm for creative results a comparison with a successful and established granular synthesis engine as reference is proposed.
In a listening test, where subjects are presented with one signal processed with the reference and another signal processed with the presented algorithm Recurring to a two alternative forced choice method, subjects are asked to make a choice of preference for creative purposes between the two signals presented.
As an example two files are presented here, one processed with *"Granulator"* plug in for Ableton Live, and other processed with the proposed algorithm. In one case no spatial processing or envelope have been applied to the output. In the other case spatial processing has been applied on both engines. (note that the exact control over the parameters used in the "*granulator*" for Ableton are not clear). Audio files are: "*vibra_ableton_spat.wav*"; "*vibra_ableton_nospat.wav*"; "*grainspat_spat.wav*" and "*grainspat_nospat.wav*"

4. ACKNOWLEDGMENTS

I would like to thank the lecturers W. Martens and E. Manor of the Graduate Program in Audio and Acoustics of the Faculty of Architecture, Design and Planning, of the University of Sydney.

5. REFERENCES

Roads, C. (1988). Introduction to Granular Synthesis. Computer Music Journal, Vol.12, No. 2, pp.11-13

R. Manson and F. Rumsey, (February, 2002) An Assessment of the Spatial Performance of Home Theatre Algorithms. In AES 108^{th} Convention, Paris, France