# Derivation of SOM-G Granular Synthesis Instruments from Audio Signals by Atomic Decomposition

**Paulo R. G. da Silva**

pgonsilva@gmail.com

Abstract. *The derivation of SOM-G granular synthesis instruments from recorded sounds by an analysis system based on the matching pursuit algorithm is presented. The implementation of the matching-pursuit algorithm and the structure of the dictionary of Gabor atoms are discussed. Audio signals recorded from acoustical musical instruments are analysed and compared with the reconstructed signals.*

## 1. An Analysis and Synthesis Experiment in Granular Synthesis

This paper presents the implementation of a system of analysis and synthesis of audio signals based on an atomic model of signal. The decomposition of audio signals into Gabor atoms is done by an implementation of the matching-pursuit algorithm. The atoms that result from the decomposition of an input signal are coded as an instrument in SOM-G language, that can be rendered into an audio signal and allowing for comparison between original and reconstructed signals.

The main objective of the implementation of the matching pursuit algorithm described in this paper is to obtain decompositions of a signal over a dictionary of Gabor atoms which duration is less than 100ms. Such durations are appropriate to the SOM-G instruments definition as will be discussed. The modelling of the sound of acoustical musical instruments as SOM-G instruments was the motivation for this implementation. So, obtain a compact decomposition of a signal is desirable, but was not the main objective of the implementation. The implementation can handle signals of different time-frequency characteristics. This is required for the decomposition of recorded samples from acoustical instruments that usually have transients and almost stationary parts in the same signal.

## 2. Introduction

The physicist Denis Gabor stated that a signal could be represented by a linear combination of elementary signals, named atoms or acoustical quanta [Gabor 1946]. He proposed a signal model in which time-domain and frequency-domain information are not dissociated, and suggested that the expansion in terms of atoms was more meaningful than Fourier analysis because the signals was considered simultaneously in time and frequency domains [Gabor 1947].

The model of Gabor inspired the synthesis technique named Granular synthesis, in which a signal is composed by a large number of short duration sounds named grains or atoms [Roads 1988]. Xenakis was the first to explain a compositional theory for granular synthesis [Xenakis 1963]. He proposes a possible approach to the model of Gabor in the context of an analog synthesis implementation, using sinusoidal waves of around 40 ms of duration modulated by rectangular envelopes. Curtis Roads systematically researched granular synthesis between 1975 and 1981, and is responsible for the first effective implementation of the technique [Roads 1987], [Roads 1988]. Barry Truax made the first real time granular synthesis experiment using a digital signal processing hardware [Truax 1988]. The difficulties on the generation and regulation of grains in granular synthesis has been evidenced since the first implementations, as it is usually necessary hundreds or thousands of grains per second to produce granular events. The active research on granular synthesis in the last years brought up various approaches to grains generation and regulation, and granular synthesis was used to create entirely new sounds. Several new approaches were developed. Some few examples show the variety of new approaches to granular synthesis regulation: cellular automata as granular regulation mechanism [Miranda 1995], granulation and synthesis from natural sounds as granular generation, allowing time or pitch transformations [Jones and Parks 1988],[Truax 1994], [Keller and Truax 1998], applications of group theory to granular synthesis [Fabbri and Maia Jr 2007], among other works.

Analysis-synthesis systems provide a conceptual framework for the development of signal modelling methods and their applications. The existence of a feasible analysis method for granular synthesis allows that the analysed signal be compared with the reconstructed signal so that the atomic model and the implementation can be tested.

There are some analysis methods that can derive time-frequency signal models. The Wavelet transform can be used to extract time-frequency information from audio signals [Kronland-Martinet 1988],[Faria 1997]. Basis pursuit applies modern linear algebra techniques to decompose a signal into an optimal combination of atoms chosen from a base [Chen, Donoho, and Saunders 1998]. Matching-pursuit [Mallat and Zhang 1993] is a greedy algorithm for the atomic decomposition in terms of atoms chosen from a dictionary.

The matching-pursuit algorithm is the analysis method that was implemented in the system described in this paper because its simplicity, stability and flexibility. Some improvements on the performance of the original algorithm has been presented, like Fast Matching Pursuit [Gribonval 2001] and Harmonic Matching Pursuit [Gribonval, Bacry 2003]. Improvements on the resolution of the analysis were brought by High Resolution Matching Pursuit [Gribonval, Bacry, Mallat, Depalle, Rodet 1996], and a measure of the destructive interference between atoms can be found in [Shynk, Daudet and Roads 2008].

## 3. Gabor Atoms

The greatest part of the theory of communication of the early twentieth century was developed on the basis of Fourier theorem. According to Gabor, though the Fourier method is mathematically correct, the physical interpretation of the results is somewhat difficult to reconcile with physical intuitions [Gabor 1946]. For human hearing, time and frequency patterns are associated in sound perception, but in Fourier theory time and frequency domains are mutually exclusive.

Gabor proposed a signal representation that reveals both its time and frequency structures. All the mathematical development can be found in [Gabor 1946] and [Gabor 1947], and we will just highlight the main results. The time frequency localization of each atom is constrained by a resolution limitation similar to the Heisenberg uncertainty principle of quantum mechanics.

$$\Delta t \Delta f \geq 1 \quad (1)$$

The inequality in (1) establishes an important relation between time and frequency resolution. In order to achieve the best time and frequency discrimination, the ideal form of the elementary signals should be one for which the product $\Delta t \Delta f$ has its minimal value and the inequality (1) becomes an equality. The signal for which $\Delta t \Delta f$ is unitary is the product of a harmonic oscillation by a Gaussian pulse.

$$\psi(t)=e^{-\alpha(t-t_0)^2}e^{i(2\pi*f_0*(t-t_0))} \quad (2)$$

The parameter fo is the mean frequency of the atom, and to is the mean epoch. The parameter α is related to the dilation of the pulse that modulates the harmonic oscillation, and determines the effective duration of the atom and its effective frequency bandwidth.

$$\Delta t=\frac{\sqrt{\pi}}{\alpha} \quad (3)$$

$$\Delta f=\frac{\alpha}{\sqrt{\pi}} \quad (4)$$

Real atoms must have an additional parameter, the phase shift φ of the harmonic oscillation. The mathematical form of real Gabor atoms is shown by expression (5).

$$\psi_r(t)=e^{-\alpha(t-t_0)^2}\cos(2\pi*f_0(t-t_0)+\varphi) \quad (5)$$

Figure 1 shows the aspect of a real Gabor atom for α = 20, fo=110 and φ=0. This value of α implies in Δt=88,2 milliseconds and Δf=11,28 Hertz. The dotted line represents the gaussian function that modulates the harmonic oscilation.

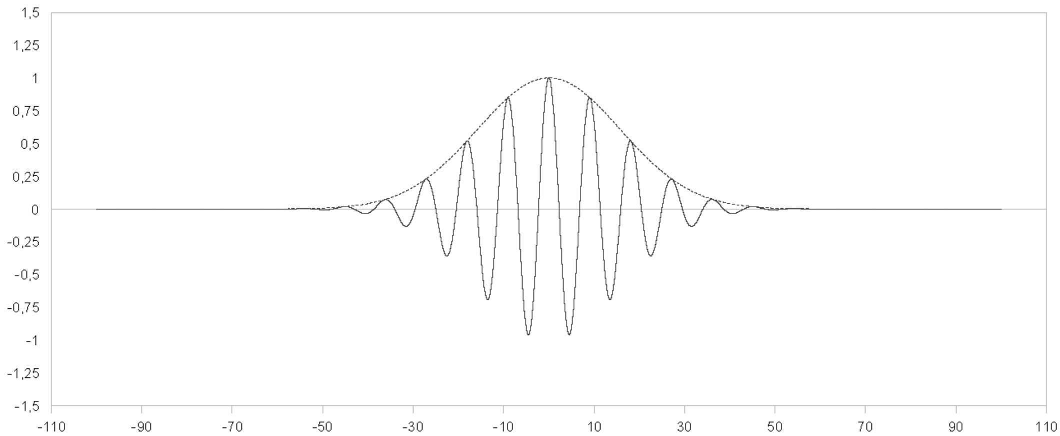


**Figure 1 – A real Gabor atom.**

Each atom can be represented as a rectangle in a time x frequency diagram . The center of the rectangle stays at the coordinates of the mean epoch and mean frequency; its width is proportional to its effective duration Δt and its height is proportional to its effective bandwidth Δf. Such diagram is called an information diagram, and the rectangles that represent atoms in an information diagram are called characteristic cells.

Figure 2 shows an information diagram and the representation of atoms as characteristic cells. The information diagram contains information about both time and frequency structures of a signal.

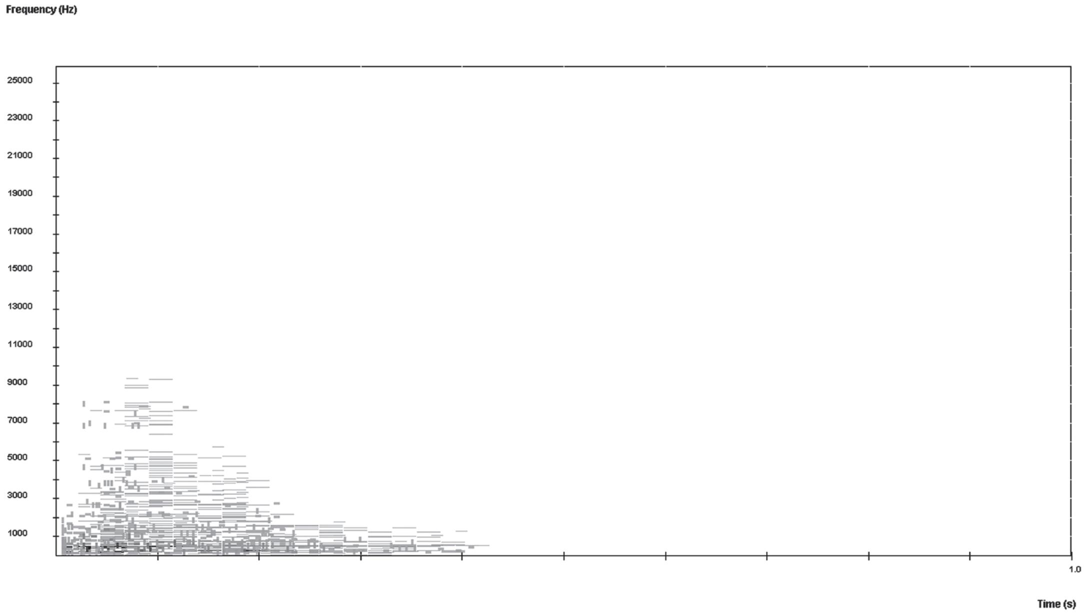


**Figure 2 – The Information Diagram**

## 4. Overview of the Matching Pursuit Algorithm

Matching Pursuit [Mallat and Zhang 1993] is a greedy iterative algorithm for deriving signal decompositions in terms of expansion functions chosen from a dictionary of basis functions or *atoms*. At each iteration, the algorithm looks in the dictionary for the atom that best approximates the signal, where the two-norm is used as the approximation metric. The contribution of the chosen atom is then subtracted from the signal and the algorithm restarts to one more iteration over the residual, until some halting criterion is met, as a residual energy threshold. The mathematical development of the algorithm and the proof of its convergence can be found in [Mallat and Zhang 1993], and a comparison with other atomic decomposition methods can be found in [Goodwin 1997].

Let D be a dictionary of complex atoms. Each function $d_k \in D$ can be characterized by its effective duration δ, its mean epoch τ and its mean frequency f. Let all atoms in D be normalized

$$\langle d_k, d_k \rangle = 1, \forall d_k \in D \quad (7)$$

The task at the i-th iteration of the algorithm is to find the atom $d_k \in D$ that minimizes the two-norm of the residual signal $r_i$. It can be shown that this is equivalent to choosing the atom whose inner product with the signal has the largest magnitude

$$d_i = arg\ max_{d_i \in D} |\langle d_i, r_i \rangle| \quad (8)$$

The i-th expansion coefficient $\alpha_i$ is the inner product between the chosen atom $d_i$ and the residual signal $r_i$.

$$\alpha_i = \langle d_i, r_i \rangle \quad (9)$$

At the end of the iteration, the term $\alpha_i d_i$ is subtracted from the residual $r_i$

$$r_{i+1} = r_i - \alpha_i d_i \quad (10)$$

After I iterations, the signal S can be represented by the expression

$$S=\sum_{i=1}^{I} \alpha_i d_i + r_{I+1} \quad (11)$$

The mean-squared error of the reconstructed signal decreases as the number of iterations increase, so matching pursuit can derive a reasonable approximation for a signal. It is well-known that matching-pursuit does not lead to optimal approximations, but greedy approaches are justified given the complexity of finding an optimal approximation, a NP-Hard problem [Goodwin 1997].

With a dictionary of Gabor atoms, a matching pursuit defines a time-frequency transform. An appropriate dicionary is required to achieve compactness, but there is a compromise between the number of atoms present in a dictionary and the number of computations necessary to choose the atom that best fits the signal at each iteraction.

## 5. An Implementation of the Matching-Pursuit Algorithm

The matching-pursuit algorithm was implemented as a java package and integrated to the implementation of the SOM-G language packages. The result of the decomposition of an audio file is expressed as a SOM-G instrument. A granular analysis/synthesis system was implemented; the SOM-G interpreter can reconstruct the signal from the granular synthesis instrument obtained . Figure 3 shows a fluxogram for the decomposition of a signal.

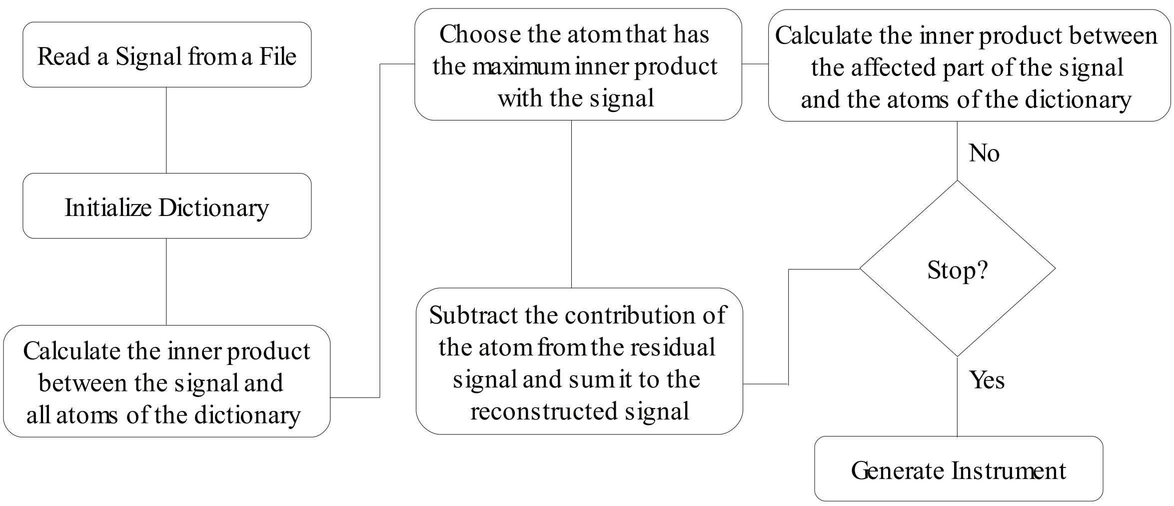


**Figure 3 – Fluxogram of the decomposition process**

In order to obtain an analytic signal from a real signal, a Hilbert transform is done over it. This is not a requirement of the matching-pursuit algorithm that can be implemented with real atoms by the introduction of a phase parameter in the dictionary, but complex atoms incorporates the phase as an implicit parameter and lead to a simpler algorithm. After the decomposition, the atoms can be converted to real signals and the phase can be extracted from the complex coefficients that result from the decomposition.

The computation of the correlations $\langle d_i, r_i \rangle$ for all $d_k \in D$ is costly, so the implementation previewed a strategy to avoid unnecessary computations. The atoms used in the implementation are finite, and each atom extracted from the residual signal affects only part of the signal. At each iteration the correlations are stored, and when the atom that has the largest magnitude of correlation is chosen, only the correlations in the part affected by the subtraction of its contribution must be recalculated for the next iteration. The class diagram of the package atomic_decomposition is shown in figure 4.

The dictionary composed only by Gabor atoms was constructed with only five effective durations for most of the signals that were decomposed: 3, 6, 12, 24 and 48 milliseconds. For each duration, the frequencies are distributed according to the interval calculated by the relation (1), from a minimal fixed value to half of the sampling rate of the analysed signal, according to Nyquist sampling theorem. The translation of the atoms are fixed as the effective duration of the atoms.

The diagram of the classes in the package atomic_decomposition is shown in figure 5. The package has only three classes: AtomicDecomposer, GaborDictionary and Signal.

The class *AtomicDecomposer* implements the matching pursuit algorithm. It has a constructor that accepts as argument a reference for an audio file. The code bellow shows the creation of an instance of the AtomicDecomposer class:

```
mp = new AtomicDecomposer(new File("sample.wav"));

mp.start();
```

The class *GaborDictionary* has its structure defined by an array that stores the durations in milliseconds of the grains:

```
durations[0]  = 0.003f;
durations[1]  = 0.006f;
durations[2]  = 0.012f;
durations[3]  = 0.024f;
durations[4]  = 0.048f;
```

A new instance of the GaborDictionary class can be created as follows.

```
/* Creates a Gabor Dictionary with minimum frequency of 15
Hz, maximum frequency of 44100 Hz and sample rate of 44100
Hz */
DC = new GaborDictionary(15, 22050, 44100);
```

The class *Signal* can represent a complex signal of one or two channels, and has many convenience methods. An example of the creation and initialization of an instance of the Signal class is shown bellow:

```
sg = new Signal(new File("sample.wav"));

sg.create_analytic_signal();
```

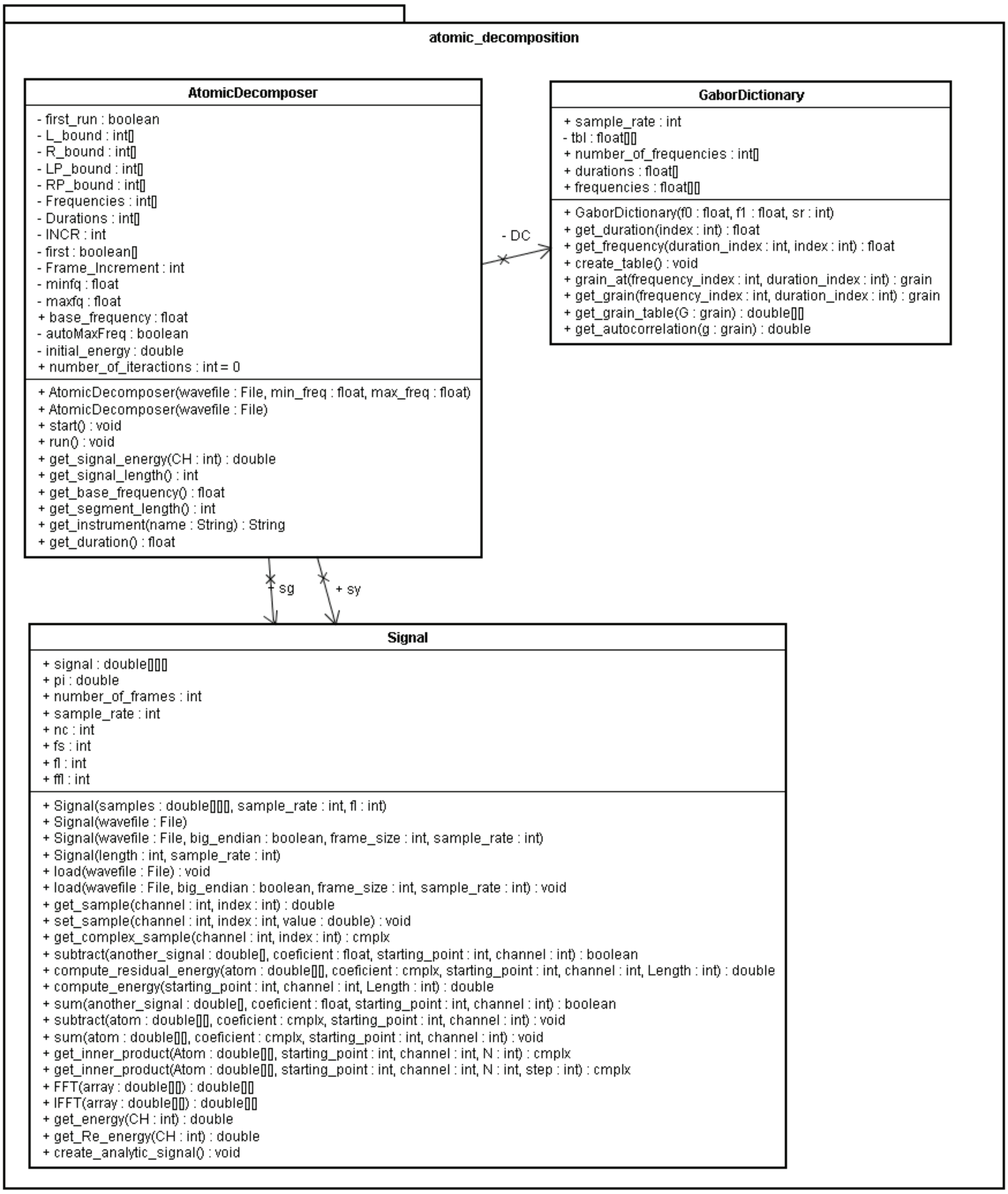


**Figure 4 – Package atomic_decomposition – Class Diagram**

## 6. Results

The decomposition and resynthesis of a berimbau note is shown bellow. A berimbau is an african percussion instrument. It has only one string, that is played with a wood stick and a rock.

Figure 5 shows the recorded signal. Figure 6 shows the reconstructed signal. Figure 7 shows the spectrum of the analysed signal, and figure 8 shows the spectrum of the resynthesized signal. The signal was recorded at 44100 Hz, 16 bits. The analysis resulted in 6965 grains for each channel.

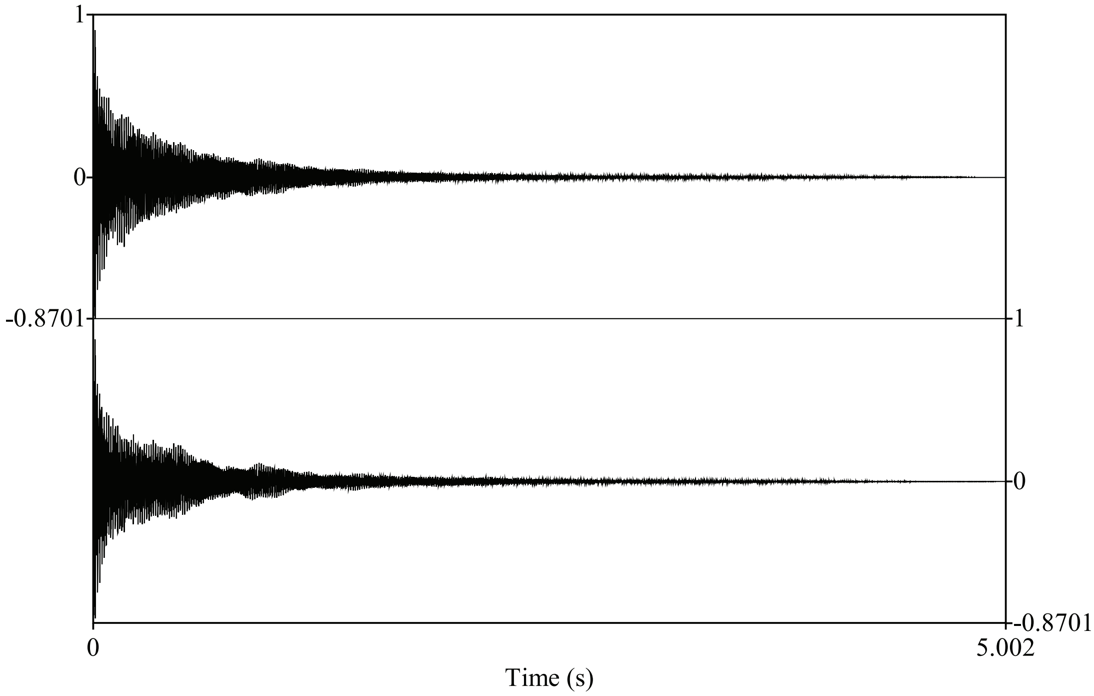


**Figure 5 – The original recorded signal**

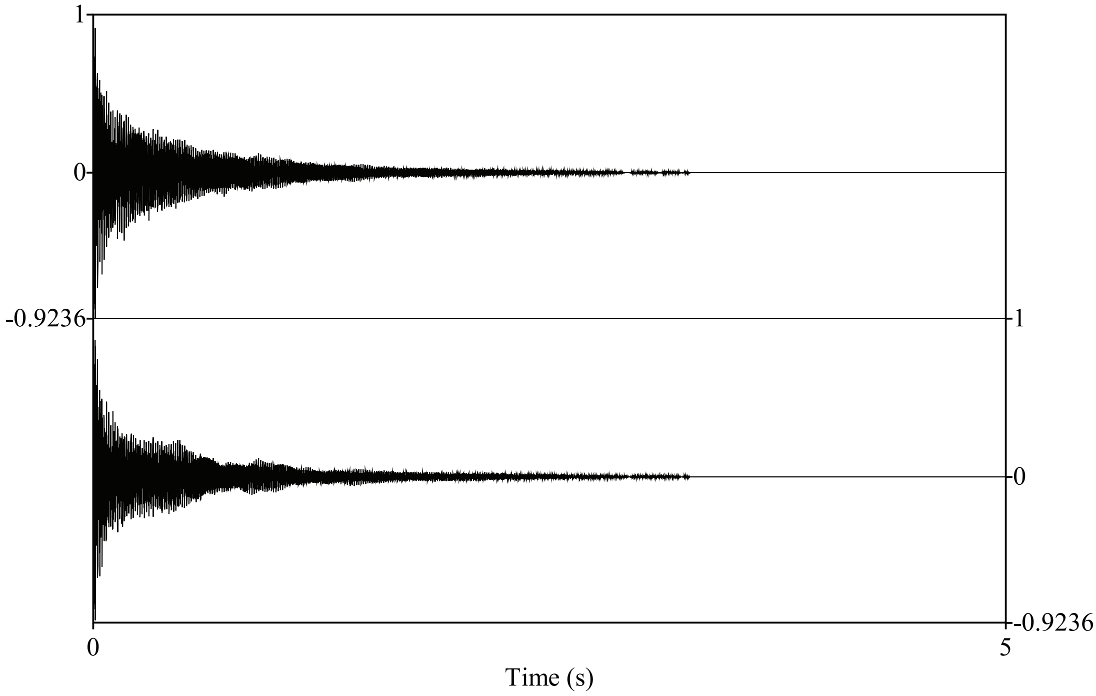


**Figure 6 – The resynthesized signal**

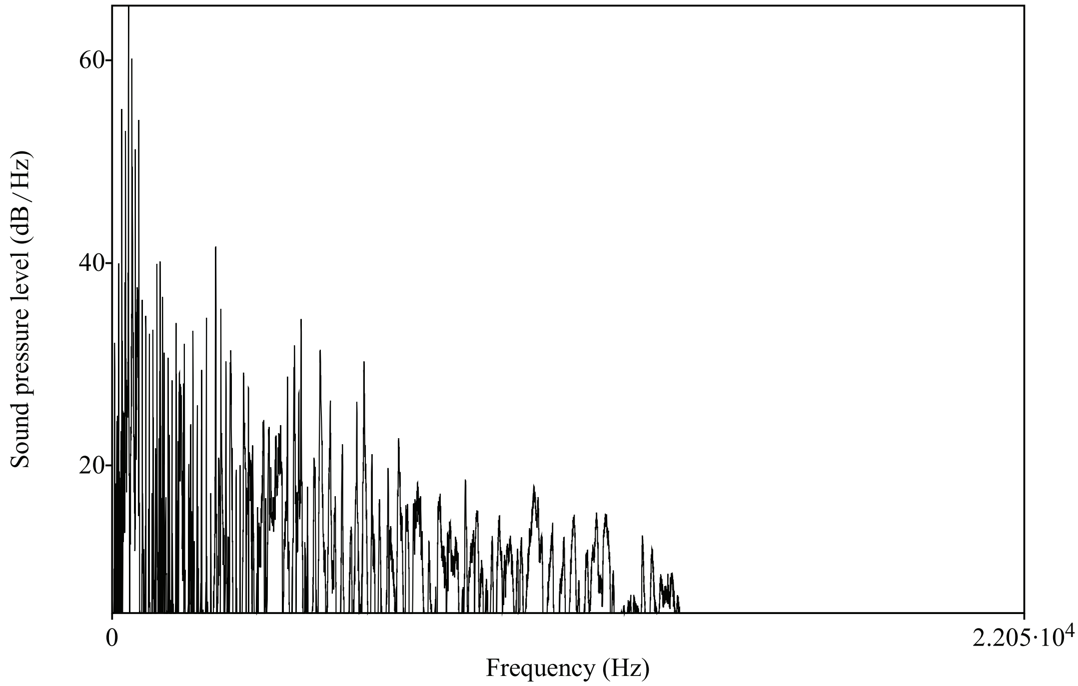


**Figure 7 – The spectrum of the recorded signal**

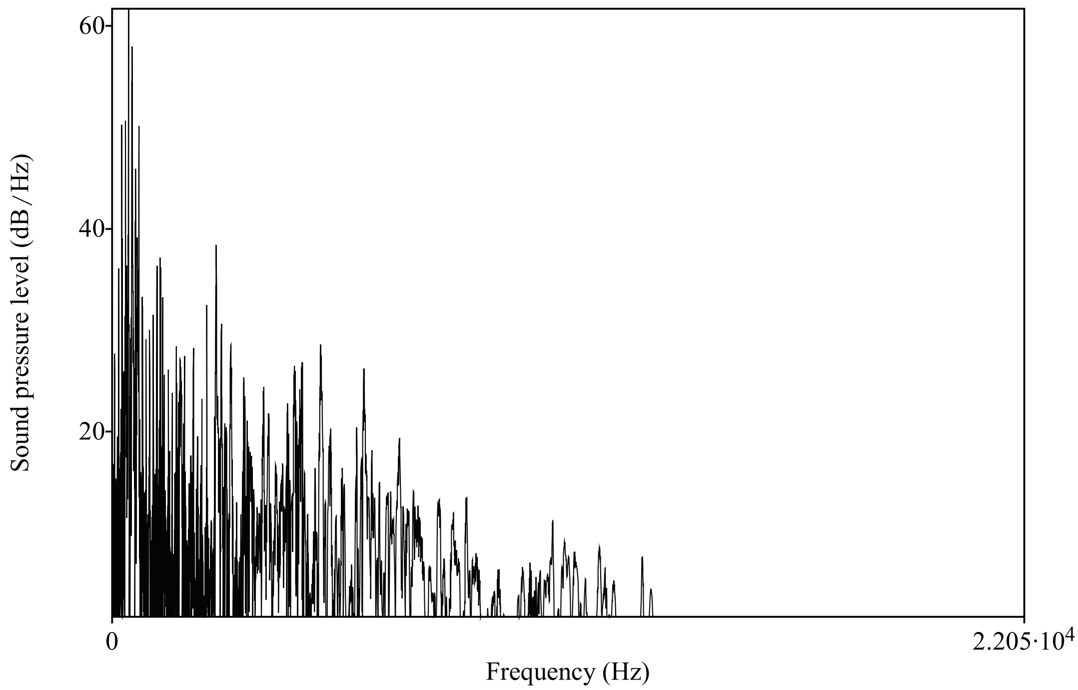


**Figure 8 – The spectrum of the resynthesized signal**

## 7. Future Work

A bank of granular synthesis instruments derived from acoustical instruments can be constructed and employed for music composition applications, improving the musical possibilities of the SOM-G language. A bank of phonems can also be modeled as granular synthesis instruments and applied to the design of speech synthesis systems.

An implementation of the whole system in a faster, non-interpreted language is desirable.

## 8. References

Chen, S. Donoho, D. and Saunders, M. "Atomic Decomposition by Basis Pursuit," SIAM, vol. 20, no. 1, pp. 33–61, 1998.

Fabbri, R. and Maia Jr, A. "Applications of Group Theory on Granular Synthesis". Annals of the VIII Brazilian Symposium on Computer Music, 109-120, 2007.

Faria, R. R. A. "Aplicação de Wavelets na Análise de Gestos Musicais em Timbres de Instrumentos Acústicos Tradicionais" Msc thesis, Universidade de São Paulo,1997.

Gabor, D. "Theory of Communication". J. Inst. Elec. Eng. (London) 93,429-457, 1946.

Gabor, D. "Acoustical Quanta and the Theory of Hearing". Nature 4044,591-594,1947.

Gonçalves, P. and Arcela, A. "SOM-G, a Language for Granular Synthesis". Annals of the VIII Brazilian Symposium on Computer Music, 33-43, 2001.

Goodwin, M. "Adaptive Signal Models: Theory, Algorithms and Audio Applications". PhD thesis, University of California, Berkeley. 1997.

Gribonval, R., Bacry, E., Mallat,S. , Depalle, Ph. , Rodet, X. "Analysis Of Sound Signals With High Resolution Matching Pursuit ". Proc. of IEEE TFTS, 125-128. 1996.

Gribonval, R. "Approximations Non-Linéaires pour l'Analyse des Signaux Sonores". PhD thesis, Université de Paris IX Dauphine, Paris. 1999.

Gribonval,R. "Fast Matching Pursuit with a Multiscale Dictionary of Gaussian Chirps". Signal Processing, IEEE Transactions on Volume 49, Issue 5, Page(s):994 – 1001. May 2001.

Gribonval, R.; Bacry, E. "Harmonic Decompositions of Audio Signals with Matching Pursuit". Signal Processing, IEEE Transactions on, Volume 51, Issue 1:101-111. Jan. 2003.

Jones, Douglas L. e Parks, Thomas W. "Generation and Combination of Grains for Music Synthesis". Computer Music Journal, vol. 12, No. 2,27-34, 1988.

Keller, D., & Truax, B. "Ecologically-based Granular Synthesis", Proceedings of the International Computer Music Conference. Ann Arbor, MI:University of Michigan. 1998.

Kronland-Martinet, R. "The Wavelet Transform for Analysis, Synthesis, and Processing of Speech and Music Sounds". Computer Music Journal 12(4), MIT Press. 1988.

Mallat, S. and Zhang, Z. "Matching Pursuit with Time-Frequency Dictionaries". IEEE-SP, 41(12):3397-3415,1993.

Miranda, E. R., "Granular Synthesis of Sounds by Means of a Cellular Automaton", Leonardo, Vol. 28, No. 4, 1995.

Roads,C. “Introduction to Granular Synthesis”. Computer Music Journal, 12(2):27-34, 1988.

Roads, C. “Granular Synthesis of Sound”. In Roads, C., Foundations of Computer Music, Cambridge, Massachussets: MIT Press. 1987.

Truax, B. “Real-Time Granular Synthesis with a Digital Signal Processor”, Computer Music Journal, 12(2), 14-26 1988.

Truax, B. “Discovering Inner Complexity: Time Shifting and Transposition with a Real-Time Granulation Technique”, Computer Music Journal, 18(2), 38-48. 1994.

Xenakis, I. “Musiques Formelles”. La revue musicale, double numéro 253 et 254. Paris, France: Éditions Richard-Masse, 1963.

Jones, D.L. and Parks, T.W. “Generation and Combination of Grains for Music Synthesis”. Computer Music Journal, vol. 12, No. 2,27-34, 1988.

# 12°

# Simpósio Brasileiro de Computação Musical

12th Brazilian Symposium on Computer Music

# MUSIC PAPERS