

# **Sonogram ± An Acoustic Signal Analyzer / Editor**

*Version 0.90(Beta)*

by

Hiroshi Momose

Department of Zoology, University of California  
Davis, CA 95616, USA  
hmomose@ucdavis.edu

## **Introduction**

This program allows you to analyze time-frequency characteristics of non-stationary signals. It displays "sonogram" ± two dimensional density plot of sound energy distribution. X-axis shows time running from left to right, and Y-axis shows frequency (upper in the graph, higher the sound ). It allows you to "see" how a signal sounds like.

## **Analysis method**

The program uses one of the several time window functions to cut out a small portion of signal samples, and perform spectrum analysis on them. You can use either FFT or auto-Wigner-transform to calculate frequency spectrum, which is a vector containing the sound pressure level at each discrete frequency. The program then draws this spectrum as a column of pixels with different darkness levels. A darker pixel means a higher energy at that particular time and frequency point.

## How to display a Sonogram

To display a sonogram, double click on the program icon, open a sound file using **Window / Open** menu. A window showing the waveform envelope of the signal appears at the bottom of the screen. Select a part of the signal you want to analyze by clicking on the start point and dragging to the end point, or, if you want to select the entire signal, click **Edit / Select all** menu. Then click **New Sonogram** in the main menu. A window named "**Control Panel**" appears at the top of the screen, and then another window called "**Sonogram**" appears. It immediately starts to display the sonogram of the selected portion. If you want to interrupt this process, just click a mouse once inside the sonogram window. To re-display it, click on

the button called **Display** in the sonogram window. Each time you select **New Sonogram** menu, a new sonogram window appears. You can also open more than one sound file at the same time. If you want to close any window, click inside that window and choose **Window / Close** menu, or just click the close button in the upper right corner of each window. You can't close the control panel, you can just hide it.

## Measuring the time and frequency of the signal

You can click inside the sonogram view to measure and time and frequency at the cursor point. They are displayed in text cells below the view, labeled **From: Sec** and **Hz**. If you drag the cursor, the values at the cursor point are shown in **To: Sec** and **Hz** cells, and the time and frequency interval between that point and the initial point are displayed in **Δ: Sec** and **Hz**. Also, as you click or drag the cursor inside a sonogram view, the power spectrum at the cursor point is calculated and displayed real time in the Control Panel. You can then click inside this power spectrum to measure the exact frequency and sound pressure level ( in dB  $\pm$  decibel ).

## Changing the appearance of sonogram display

You can control many aspects of sonogram display using the control panel.

I will explain them item by item.

## Analysis Options

Left part of the control panel is where you control the sound analysis features.

- **Analysis method**

You can select different analysis methods by this pop-up menu. The default method is **FFT** (Fast Fourier Transform). Another item named **Wigner** selects Wigner-Ville transform. This method gives a better time-frequency resolution, but it takes longer to compute. Also, since this transform is not linear, cross term interaction products between each time/frequency component pair appear as artifact. For example, if your signal has two components at 1kHz and 2kHz, the cross term will appear at 1.5 kHz. You can select an item called **FFT & Wigner** to reduce this cross term effect. This method computes the geometric mean between FFT power spectrum and Wigner transform. In the example above, the power spectrum has no component at 1.5 kHz, so Wigner cross term at this frequency is

masked and doesn't show up. For wide band signals, things get far more complicated than this example, and Wigner transform doesn't work very well. If you want to know more about Wigner transform, refer to Boashash, 1991 for a review.

- **Window shape**

This option selects the time window function. Available functions are...

Hanning window

$$W_n[n] = \begin{cases} 0.5 - 0.5\cos(2\pi n/N) & \text{for } 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

Hamming window

$$W_n[n] = \begin{cases} 0.54 - 0.46\cos(2\pi n/N) & \text{for } 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

Blackman window

$$W_n[n] = \begin{cases} 0.42 - 0.5\cos(2\pi n/N) + 0.08 \cos(4\pi n/N) & \text{for } 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

Rectangular window

$$W_n[n] = \begin{cases} 1 & \text{for } 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

Default function is **Hanning window**. You will be happy with this in most cases. If your signal has multiple frequency components that are close together, then try using **Hamming window**.

- **Peak picking**

If you select a button inside a box named **Peak Picking**, the program will look for local frequency peaks in the power spectrum vector and displays the peak information only. The algorithm used here is based on Markel & Gray, 1976. It uses simple parabolic curve fitting to estimate the true frequency from discrete frequency values. Combination of Wigner-Ville transform and this local peak picking is ideal for tracking narrow-banded FM signals like birdsong. The cell named **SideBand** is used to delete annoying side bands. To suppress this function, type 0 and hit enter. Also, if you enter a very large value (like half the sample speed), only one maximum energy peak will be displayed at each moment.

- **Analysis Resolution / Time Window Size**

This pop-up menu selects the length of the time window. For speed, only the power of 2 values are allowed. This will affect the frequency

resolution of the analysis, which will be displayed in the cell labeled **Freq. Resolution**

- **Analysis Resolution / Time Increment**

This slider / cell enables you to change the amount of time the time window is shift each time. The default value is 128 points ( samples ). You can also enter the amount as absolute time using **Freq. Resolution** cell, but the time will be truncated to nearest discrete points. The smaller the amount of shift, the more fine-grained the analysis result will become and the longer it will take.

- **Use DSP**

If you select this button, the program will use 56001 DSP to compute FFT. By default, it uses CPU. This affects Wigner transform too, since Wigner algorithm uses FFT three time in each iteration. (In this version, this button is disabled.)

- **Sample Freq.**

This cell shows the sampling frequency. This is set automatically, but you can also change this value in case original sound was played

half-speed etc.

- **Low Cut Freq.**

This cell can be used to enter the low cut-off frequency. In Wigner-Ville transform, this helps reducing the interaction between signal component and low-frequency noise component. So, don't forget to use this option.

## **Display Options**

Right part of the control panel is where you control the display features.

- **Grayscale**

You can change the number of different gray levels to render each pixel. The default value is **4 Grays**, which uses four generic gray levels of MegaPixel Display. Or you can use **B & W** ( Black and White ) or **Many** ( Continuous tone grayscale ± This should look great on color monitors ).

- **Pixel Size**



You can change the width(X) and height(Y) of each pixel. The default is 1. *(In this version, X is not selectable and calculated automatically from the duration of the sound and amount of time increment.)*

- **Emphasis / High shape**

If you select this button, high frequency components will become darker. This version uses a simple post-emphasis ( linear scaling ) to do this.

- **Display Frequency Range**

You can use these sliders / cells to change the lower and upper frequency of the sonogram. Lowest limit is off course zero, and highest limit is half the sampling frequency( called the Niquist frequency ). The slider stopped at each integer kHz point. If you want to use a finer setting, you can enter the values in the cells.

- **Display Dynamic Range**

These sliders / cells affect the darkness of the graph, that is, how

actual sound pressure levels are assigned to gray scale values. In this version, only linear scaling is used in this conversion. You can change the lower and upper limit of dynamic range. For example, if you choose -20dB as lower limit and 0dB as upper limit, anything below -20dB will be white and anything above 0dB will be black, and the dynamic range will be 20dB.

## Other Controls

A radio-button named **Scaling** let you select the scaling of the power spectrum display in the Control Panel. Default is **Linear** scaling. You can also select **Log** scaling. A button named **Revert** initializes all the settings in the control panel to their default values.

## Editing a sound

You can edit the sound by usual cut and paste method. Also, you can fill the selection with 0 volt with **Edit / Erase** menu, and also add silence at the end of the file using **Edit / Add Silence**, which will ask you the duration of the silence you want to add. If you want to restore the sound to its original form, use the **Window / Revert To Saved** menu.

## Credit

This program is based on EdSnd 1.4 by James Pritchett and Steven M. Boker, which, in turn, is based on SoundEditor by Lee Boynton. FFT routine ( for CPU ) was originally coded by Kevin Peterson of MIT Media Lab. Wigner transform algorithm used here was originally coded by Yuji Nishimori ( Nishimori, 1988 ). All the routines were heavily modified by me, so the original authors are not necessarily responsible for any bugs that may be found in these routines.

## Reference

- Nishimori, Y. 1988 Wigner distribution and its application ( *in Japanese* ). A thesis presented to the Department of Electrical Engineering, Faculty of Engineering, Doshisha University, Kyoto, Japan.
- Boashash, B. 1991 Time-frequency signal analysis. In: *Advances in Spectrum Analysis and Array Processing* Vol.1 pp. 418-517. Ed. by Simon Haykin, Prentice Hall, New Jersey ( ISBN 0-13-007444-6 ) .
- Markerl, J.D. & Gray, A.H.Jr. 1976 *Linear Prediction of Speech*. Springer-Verlag, Berlin & Heidelberg.

