

- *Research reports*
- *Musical works*
- *Software*

AudioSculpt

User's manual

Second Edition, April 1996

This manual can not be copied, in whole or part, without the written consent of Ircam.

This manual was written by Peter Hanappe under the supervision of Marie-Hélène Serra, and was produced under the editorial responsibility of Marc Battier, Département de la Valorisation, Ircam. Proofreading by Andrew Gerszo.

AudioSculpt

Design	Philippe Depalle, Chris Rogers
Program	Chris Rogers
Project supervision	Gerhard Eckel

SVP

Design	Philippe Depalle
Program	Philippe Depalle, Gilles Poirot, Chris Rogers, Jean Carrive

Analysis modules

Algorithms	Xavier Rodet, Philippe Depalle, Guillermo Garcia, Ernst Terhardt, Boris Doval
------------	---

Documentation

Coordination	Marie-Hélène Serra
Manuals	Marc Battier, Jean Carrive, Peter Hanape, Mario Mary, Brice Pauset, Marie-Hélène Serra

This documentation corresponds to version 1.0 or higher of AudioSculpt and to version 1.3 of SVP.

Apple Macintosh is a trade-mark of Apple Computer, Inc.

AudioSculpt is a trade-mark of Ircam.

Second Edition, April 1996

Ircam
1 place Igor-Stravinsky
F-75004 Paris
Tel .01 44 78 12 33
Fax 01 44 78 15 40
E-mail : ircam-doc@ircam.fr

IRCAM Users group

The use of this software and its documentation is restricted to members of the Ircam software users group. For any supplementary information, contact:

Département de la Valorisation
Ircam
Place Stravinsky, F-75004 Paris

Tel .01 44 78 49 62
Fax 01 44 78 48 15 40
E-mail: bousac@ircam.fr

Send comments or suggestions to the editor:

E-mail: bam@ircam.fr

Mail: Marc Battier,
Ircam, Département de la Valorisation
Place Stravinsky, F-75004 Paris

<http://www.ircam.fr/musinfo>



To see the table of contents of this manual, click on the Bookmark Button located in the Viewing section of the Adobe Acrobat Reader toolbar.

Contenu

Résumé 7

Les opérations de AudioSculpt 8

Analyse 8

Traitement 8

About this manual 10

Introduction to AudioSculpt 11

The operations of AudioSculpt 12

Analysis 12

Processing 12

Playing the soundfiles 13

About soundfiles and data files 13

Soundfiles 14

Parameter files 14

File name lengths 14

Additional documentation for AudioSculpt 14

A quick tour 15

Opening a soundfile 15

Sonogram analysis 17

The sonogram filter 21

'Process in place' or 'Process and save': Transposition 22

The analysis modules 24

The command line 26

Processing and parameter files: formant filter 27

What's next? 29

Analysis 30

General principles of the short-time spectral analysis 30

Window size 31

Window step 32

Choice of window 33

Spectral analysis 33

Fast Fourier transform (FFT) 33

Amplitude and phase spectrum 34

Windowing effects 34

FFT size 35

Linear Prediction (LPC) 36

Analysis display: sonogram or text file 37

The sonogram 37

The toolbox 38

Text file output 39

Detection of partials: peak detection and fundamental estimation 41

Peak detection 42

Fundamental estimation 43

Analysis based on perceptual criteria: masking effects 45

Processing: overview 47

Filtering 51

General principles 51

The sonogram filter 53

The surface tool 54

The pen plus and pen minus 54

Band filtering 55

Breakpoint filtering 56

Formant filtering 58

Surface filtering 60

Cross synthesis 62

The cross synthesis source filter 62

Generalized cross synthesis 64

Time scale modification and transposition 67

Time scale modification 67

Constant time scale modification 68

Time-varying time scale modification 68

Improving the quality of a time-stretched sound 70

Transposition 70

Cents scale 70

Constant transposition 70

Time-varying transposition 71

Direct soundfile processing operations 73

Sample rate conversion 73

The File menu 74

New 74

Open... 75

Close 75

Save, Save As... 75

Set Default Folders 75

Page Setup... 75

Print... 75

Quit 75

The Edit menu 76

Undo 76

Cut 76

Copy 76

Paste 76

Clear 76

Select All 76

Preferences 77

Show Clipboard 77

The Processing menu 78

Time Scale Modification 78

Sample Rate Conversion 78

Transposition 78

Breakpoint Filter 79

Band Filter 79

Surface Filter 80

Formant Filter 80

Sonogram Filter 81

Cross Synthesis Source Filter 81

Generalized Cross Synthesis 81

Use Command Line 82

Run Script 82

Stop Processing 82

Sonogram Analysis 83

Realtime Analysis 84

Analysis Settings 84

Realtime Sonogram Scaling 85

Edit Color Palette 85

FFT Analysis 85

LPC Analysis 86

Peak Detection 86

Fundamental Estimation 87

Masking Effects 88

Play Sound 89

Play Selection 89

Get Sound Info 89

Sound Output 89

Show Console 89

Annex A:Command lines 91

Annex B:List of options of the command line 94

Index 100

Résumé

Ce manuel présente le logiciel AudioSculpt pour le Macintosh. AudioSculpt est un programme qui a été conçu et développé à l'Ircam. Il est basé sur SVP¹, programme créé par Gilles Poirot et Philippe Depalle, du groupe Analyse/synthèse. Sous la direction de Gerhard Eckel et avec la coordination de Marie-Hélène Serra, SVP fut ensuite révisé et optimisé par Chris Rogers, avec des extensions programmées par Jean Carrive.

Conçu à partir de la version 1.3 de SVP pour le Macintosh, AudioSculpt est une interface graphique de représentation des sons ainsi que de contrôle d'un ensemble de transformations compression/dilatation temporelle, transposition, filtrage, synthèse croisée, etc.).

La version 1.0 d'AudioSculpt intègre toutes les commandes déjà disponibles dans SVP 1.3A du Macintosh, et ajoute de nouvelles possibilités, essentiellement l'analyse spectrographique (sonogramme), ainsi que la possibilité de filtrer le son en dessinant directement sur le sonogramme. D'autres types de traitements contrôlés graphiquement seront prochainement ajoutés à cet ensemble, tel que l'édition d'une courbe de compression/dilatation temporelle ou de transposition.

AudioSculpt est utile à la fois pour les chercheurs qui travaillent sur la structure des sons et pour les musiciens qui veulent modifier des sons. Les applications sont nombreuses, car il est possible, avec AudioSculpt, de modifier indépendamment l'amplitude, la fréquence, le spectre et la durée des sons, quelque soit leur origine. AudioSculpt fournit par ailleurs des résultats d'analyse hautement sophistiquées sous la forme de fichiers utilisables par d'autres programmes. Enfin, l'utilisateur peut directement intervenir sur les zones d'affichage graphique, ce qui rend aisé la manipulation d'opérations complexes.

1. SVP : Super vocodeur de phase.

Les opérations de AudioSculpt

AudioSculpt offre deux grandes familles d'opérations : analyse et traitement.

Analyse

AudioSculpt reprend toutes les fonctionnalités précédemment disponibles dans SVP 1.3a, en ajoutant une fonction nouvelle : l'analyse spectrographique (le mode sonogramme).

L'analyse spectrographique permet le traitement des sons grâce à son interface graphique. L'utilisateur intervient directement dans la zone affichage du sonogramme obtenu.

Les autres opérations d'analyse sont :

- la Transformée de Fourier (FFT) ;
- l'analyse par prédiction linéaire (LPC) ;
- la recherche de pics dans le spectre à court terme (Peak Detection) ;
- la détection de la fondamentale (Fundamental Estimation) ;
- l'évaluation de la ou des hauteurs virtuelles (algorithme de Terhardt).

Elles constituent le module d'analyse. Le module d'analyse ne permet pas de modifier les sons : les résultats d'analyse sont recueillis dans un fichier texte ou affichés en superposition sur le sonogramme.

Un effort particulier a été fait pour la présentation des résultats d'analyse, qui sont aisément mis en forme par l'utilisateur.

Traitement

Les commandes de traitement permettent de définir des modifications à effectuer sur le son. Elles reposent sur le principe du vocodeur de phase, dans lequel le son est d'abord soumis à une analyse de son évolution spectrale dans le temps ; le traitement spécifié altère les résultats de l'analyse, et le son résultant est obtenu par resynthèse.

La plupart des traitements de son sont effectués à partir d'une analyse/synthèse par vocodeur de phase. Seuls, la transposition (sans correction) et le rééchantillonnage n'utilisent pas l'analyse/synthèse.

Dans tous les traitements à base d'analyse/synthèse, pour obtenir un résultat satisfaisant, il convient de choisir les paramètres de l'analyse (taille de fenêtre, pas d'avancement, taille de la FFT, type de fenêtre).

Parce que cette méthode garantit une indépendance des paramètres, une grande finesse de manipulation est possible. Toutefois, comme les paramètres utilisés par AudioSculpt pour l'analyse doivent être choisis avec soin, une documentation spécifique sera publiée ; elle guidera l'utilisateur dans le choix des paramètres en fonction du son à traiter et des résultats visés.

Notez que dans le menu d'AudioSculpt, les traitements (**Processing**) sont placés avant l'analyse (**Analysis**) ; la raison est que les traitements des sons disponibles possèdent leur propre mode d'analyse, tandis que les fonctions offertes par le menu **Analysis** permettent de recueillir leurs résultats soit dans un fichier, exploitable par d'autres programmes, soit par affichage graphique. Les fonctions des deux menus sont donc indépendantes.

Seul le sonogramme échappe à cette règle, puisque, sous le menu **Analysis**, il permet de définir des traitements qui seront exécutés au moyen d'une fonction du menu **Processing**.

About this manual

The manual is divided into 4 parts :

- Overview and Quick tour
- Analysis
- Processing
- Menu reference

As an introduction to AudioSculpt we will give an example of a short working session. This example is meant to explain some of the basic features of AudioSculpt and its graphical interface. If you are already familiar with AudioSculpt, you can skip this chapter.

In Part 2 you will find a more detailed description of the analysis.

In Part 3 the processing functions are explained.

In Part 4 you will find a list of all the commands in the menus.

Part 1

Overview

1 Introduction to AudioSculpt

AudioSculpt is a sound processing software package that was designed and developed at Ircam. It is based on SVP (Super Vocodeur de Phase), a program created by Gilles Poirot and Philippe Depalle of the Analysis/Synthesis team. Under the supervision of Gerhard Eckel and with the coordination of Marie-Hélène Serra, AudioSculpt has been revised and optimized by Chris Rogers, of the Interface and Sound Representation team, and other extensions were made by Jean Carrière.

Based upon SVP 1.3 for Macintosh, AudioSculpt is a graphical interface for representing sounds as well as for the control of transformations (time expansion or compression, transposition, filtering, cross synthesis, etc.). It contains all of the commands previously available in SVP 1.3 for Macintosh, and adds some new possibilities. Its main new features are the spectrographic analysis (sonogram) and the filtering by drawing polygon surfaces directly on the sonogram.

AudioSculpt is an analysis/synthesis program useful to researchers interested in the structure of sound, as well as to users who wish to study and change the properties of sounds. The applications to musical composition are manifold, because it is possible to change, independently, the amplitude, frequency and time evolution of musical sounds, whether they be natural, instrumental, or vocal. AudioSculpt also offers advanced analysis results in the form of text files that can be interpreted by other programs. Finally, the user can operate directly on the graphical representation, a procedure which simplifies complex sound manipulations.

The operations of AudioSculpt

The operations that AudioSculpt offers can be divided into two main groups: analysis and processing.

Analysis

The most significant difference of AudioSculpt with respect to SVP is the spectrographic analysis (the sonogram mode). The spectrographic analysis facilitates the processing because the properties of the sound can be observed visually. The user can work directly in the window that contains the sonogram. Further possible analyses are:

- the Fourier transform;
- the calculation of the spectrum envelope with the linear prediction technique;
- peak detection in short-time spectra;
- an estimation of the fundamental frequency;
- the evaluation of the perceptual pitches and weights (Terhardt's algorithm).

These analyses form the analysis modules: they do not allow the modification of the sound. The results are recorded either in a text file or shown on the sonogram.

Processing

The processing commands offer the possibility of defining the modifications to be applied to the sound. Most of these modifications are based on the principle of the phase vocoder in which the sound is submitted to an analysis describing the spectral evolution of the sound in time. Then the spectral representation is altered according to the requested processing and the final sound is obtained by re-synthesis. Most of the processing functions perform an analysis/synthesis using the phase vocoder. Only the transposition without time-correction and the re-sampling function do not perform an analysis/synthesis.

Note: for every processing task that performs an analysis/synthesis, it is important to choose the appropriate analysis parameters (window size, window step, FFT size, window type) in order to obtain the best results.

Because the values of the parameters of the analysis can be set independently, it is possible to tune the processing precisely. To obtain the best results, great care should be taken when specifying these parameter values.

Note: the functions in the processing menu are independent of the functions in the analysis menu. In other words, it is not necessary to perform an analysis in the **Analysis** menu in order to be able to perform processing on the processing menu. The processing operation will perform the analysis by itself.

Caution: the sole exception is the sonogram filter for which the sonogram analysis should be run first.

The function of the analysis modules is to provide a better understanding of the sound. The analysis results can help for further processing or can be used in other applications such as PatchWork (a Lisp-based computer aided composition package also available from Ircam's Forum).

Playing the soundfiles

If you use a Macintosh equipped with a 16-bit audio system (like the Power Macintosh series) you can use your internal Digital-to-Analog conversion hardware, which will play 16-bit stereo soundfiles. AudioSculpt will play your soundfiles directly through this hardware, with CD quality. Make sure that the Sound Manager in your Macintosh system is version 3 or higher. Version 3 comes with System 7.5. If you use an older system, you will still be able to install Apple Sound Manager 3.0.

If your machine is equipped with 8-bit audio hardware, you should use a sound processing board such as Digidesign's Audiomedia I and II, Sound Tools I and II, ProTools, or any other 16-bit board. AudioSculpt will still be able to play your soundfiles through the Macintosh hardware, but sound quality will be poor. Newer Macintoshes with PCI bus may play directly through the Sound Manager. Third-party boards for the PCI bus (such as Audiomedia II or ProTools by Digidesign) should be available in 1996.

About soundfiles and data files

When you first use AudioSculpt, you should set the Default folders. In the **File** menu you will find an item called **Set Default Folders**. In this item choose **Sound Folder**. The dialog that comes up lets you walk through the folder structure of your computer. The Default Sound Folder is the folder where AudioSculpt looks for the soundfile when you use the command line. It should be set to the folder where your soundfiles reside. In this example, make sure that you are in the AudioSculpt 1.0 folder and then click the **Set Default Folder** button. Do the same for the parameter folder

Soundfiles

AudioSculpt processes 16-bit soundfiles and the 32-bit AudioSculpt format. If you inadvertently use an 8-bit soundfile, AudioSculpt will not proceed.

AudioSculpt is capable of processing stereo files. Be aware that stereo file processing is slower and takes up more memory than mono soundfile processing.

Three sound formats are recognized by AudioSculpt: AIFF (Audio Interchangeable File Format), the Sound Designer II format and its own AudioSculpt format. Most digital audio applications let you save the soundfile in one of these formats. The AudioSculpt format saves the samples as a floating point number using 32 bits per sample.

Parameter files

Parameter files are the files that control the parameters of a time-varying processing such as transposition, etc... All parameter files should be edited and saved as text files, i.e. plain ASCII files. The files can be edited with the internal editor of AudioSculpt (for files smaller than 20000 characters) or with any third-party text editor (for bigger files).

File name lengths

The Macintosh operating system does not allow you to use names longer than 31 characters. As AudioSculpt adds extensions to the names you choose, such as 'out' or 'console', use short file names, in any case no longer than 23 characters.

Additional documentation for AudioSculpt

As of this writing, there is currently one other manual available for AudioSculpt : *Cross-Synthesis Handbook* by Mario Mary. This cross-synthesis tutorial comes with a DAT cassette of examples, available from Ircam's Forum. Also, make sure to consult the latest issue of AudioSculpt Newsletter for up-to-date information.

So, let's start. When you've installed AudioSculpt on your hard disk as explained in AudioSculpt newsletter, you should find two items in the AudioSculpt Folder, namely AudioSculpt 1.0, which is the application, and a file called `english.msg`. Two soundfile are available separately: `Clarinet.AIFF` and `Cymbal.AIFF`. They are distributed with this manual. You can also obtain them from Ircam's Forum. Now launch the application by double-clicking its icon. The menu bar changes and should look like this:

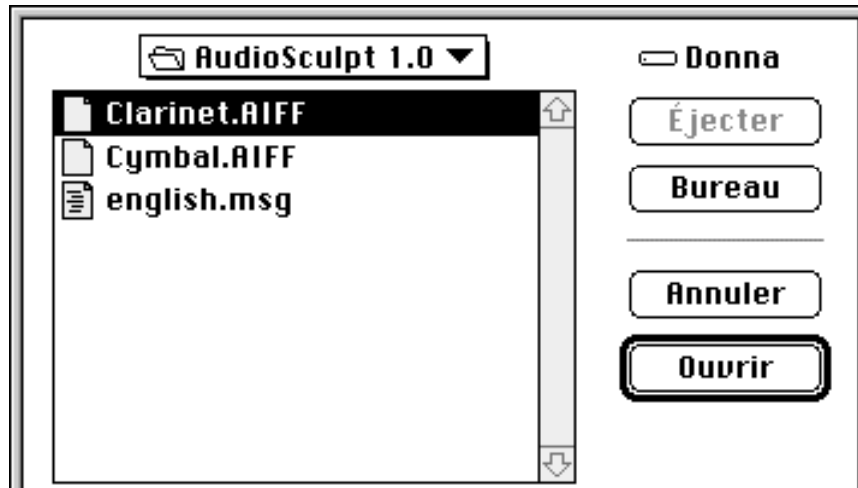


You will find all functions and commands that AudioSculpt offers in these menus. If the name of a command is shown in grey, the command is deactivated.

Opening a soundfile

To open a soundfile, choose **Open** (**Ouvrir** in French operating system) in the **File** (**Fichier** in French) menu and you will see the familiar file dialog¹:

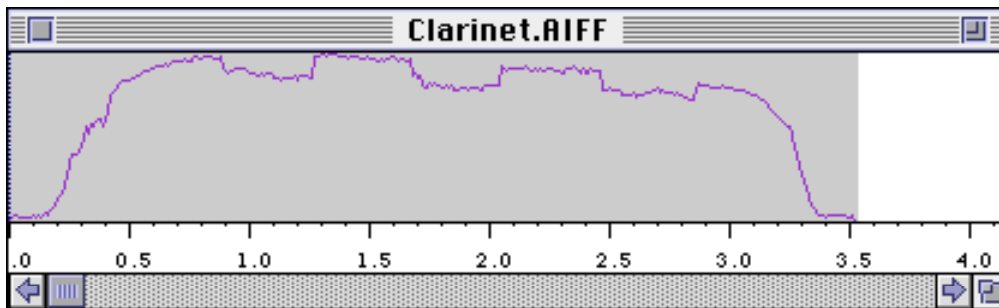
1. All standard Macintosh dialog boxes appear in French in this manual, as all examples were made on a French Macintosh system.



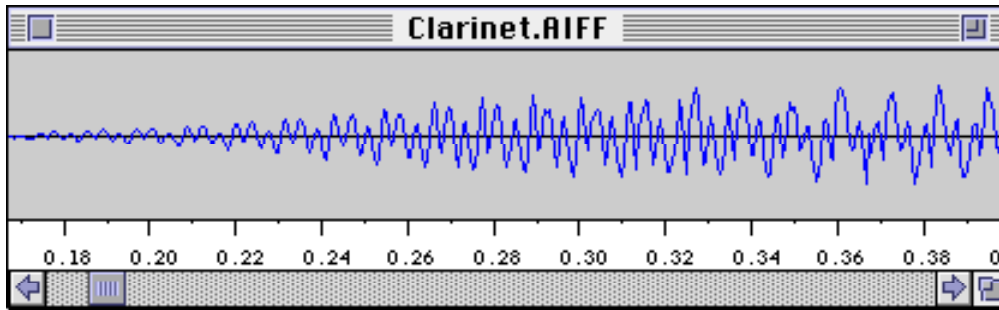
Open the soundfile **Clarinet.AIFF** by selecting it and then clicking the **Open** button (**Ouvrir** in French). The soundfile will appear in a new window on the screen. It is represented as a time-amplitude envelope curve:

- the horizontal axis shows the time in seconds,
- the vertical axis shows the amplitude envelope of the sound.

Note that AudioSculpt displays the positive amplitude envelope.



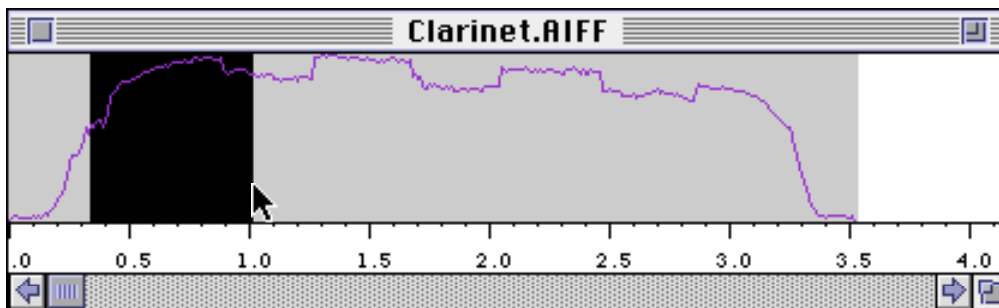
Click with the mouse in the time axis bar, hold down the mouse key and move the mouse to the right. The time axis re-scales while you're moving and the sound pane stretches simultaneously. Keep stretching the time axis until the envelope curve disappears and changes into the actual sampled sound curve. Depending on the scale, you will see the amplitude envelope or the sampled sound curve. Now double-click on the time axis to bring it back to its original scale. There is also another way of changing the time scale. Hold down the shift key and select a region in the time axis with the mouse. Release the mouse key, and you will see that the selected region is stretched to fit in the window size. Double-clicking the axis brings it back to its initial size.



The signal window behaves like any standard window: you can resize it, move it on the screen, and the scroll bar lets you display the hidden part of the sound.

Choosing **Play Sound** in the **Sound** menu will let you hear the sound. A quicker way to do this is to press the space bar. Pressing it again will stop the sound from playing.

Let's select a part of the sound. To do so, click in the sound pane and mark the region with the mouse while holding the mouse key down. The selected region is now highlighted in black.



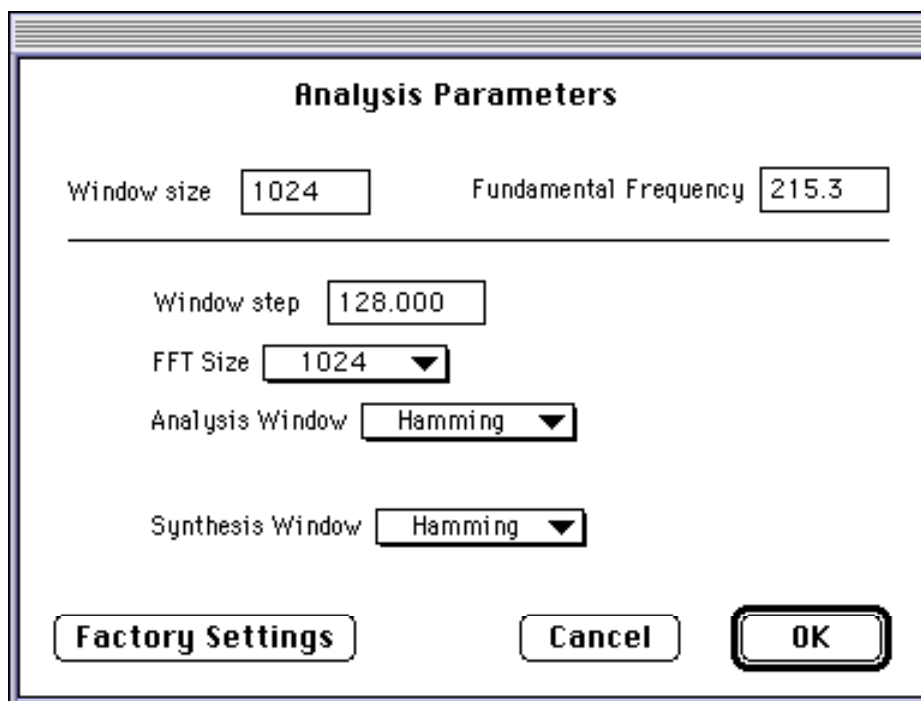
Press the space bar and you'll notice that you only hear the selected sound. Pressing the command key (the clover/apple key) and the spacebar at the same time will let you hear the complete sound again.

Sonogram analysis

Now that you can get around with soundfiles, let us analyze them. In general AudioSculpt works with the sound in the active window. You know the window is active when its title bar has its normal appearance.

Choose **Sonogram Analysis** in the **Analysis** menu. A dialog box will appear and ask for the analysis parameters. The default settings will do for now, so click the button **Factory Settings** and then click **OK** to start the analysis. Two new panes are added to the original window and then the sonogram is drawn gradu-

ally.

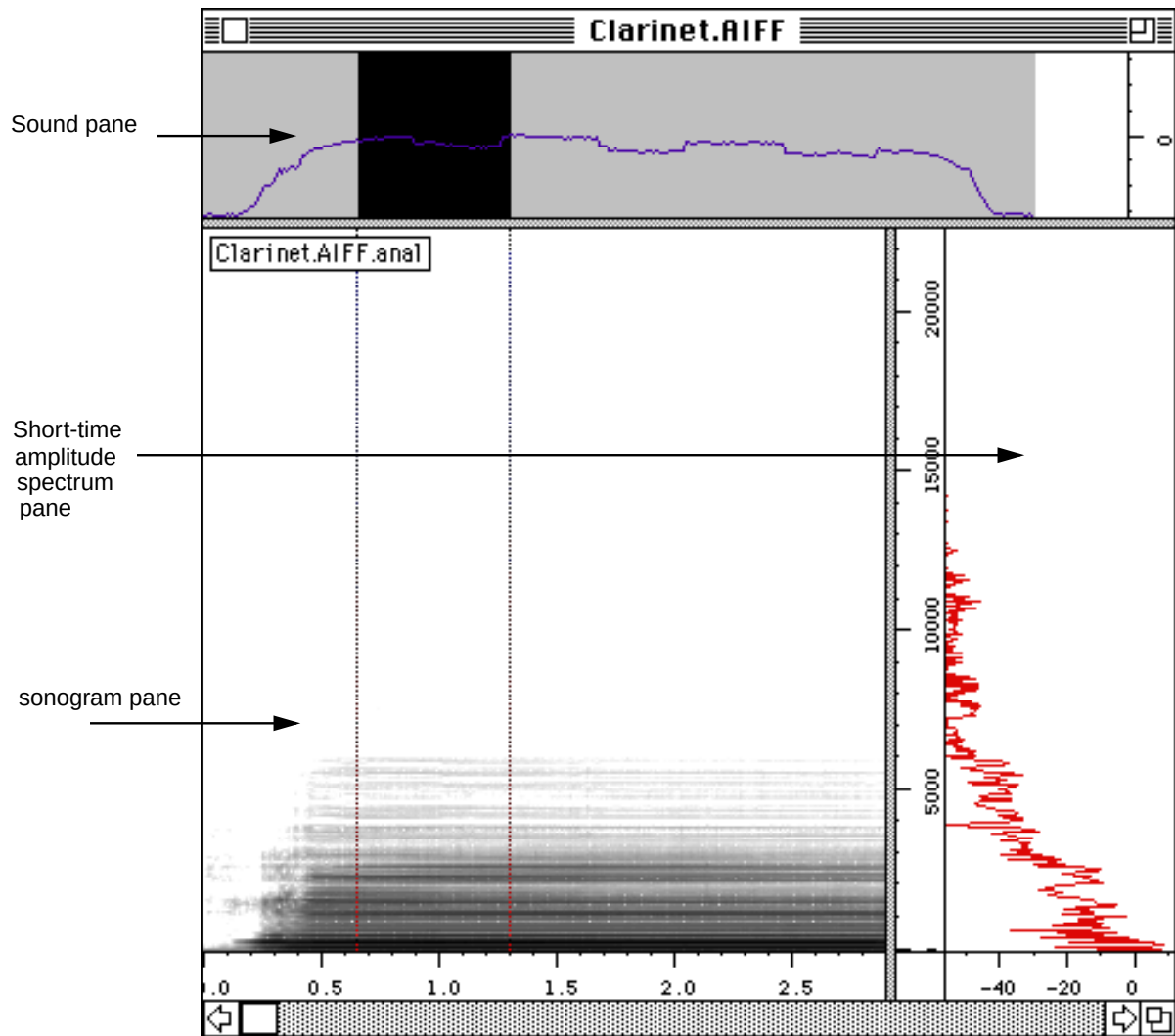


The image shows a dialog box titled "Analysis Parameters". It contains several input fields and dropdown menus. At the top, "Window size" is set to 1024 and "Fundamental Frequency" is set to 215.3. Below these, "Window step" is 128.000. "FFT Size" is a dropdown menu currently showing 1024. "Analysis Window" and "Synthesis Window" are also dropdown menus, both currently showing "Hamming". At the bottom of the dialog are three buttons: "Factory Settings", "Cancel", and "OK".

The complete sound is analyzed even if you have selected only a small part. When the analysis is finished you should see something like the figure below on your screen. The window now contains three separate panes.

- the upper pane is the **sound pane** as before, except that it has no longer a time axis.
- in the lower left hand corner of the window is the **sonogram pane**. The time axis (in second) is shown horizontally at the bottom of the pane. The vertical axis on the right side of this pane is the frequency axis (in Hertz). The sound pane and the sonogram pane share the same time axis.
- The lower right hand pane is the **short-time amplitude spectrum pane**. It has a frequency axis which it shares with the sonogram pane and an amplitude axis on the bottom of the pane. Each of these axis (time, frequency or amplitude) can be rescaled as we've done previously.

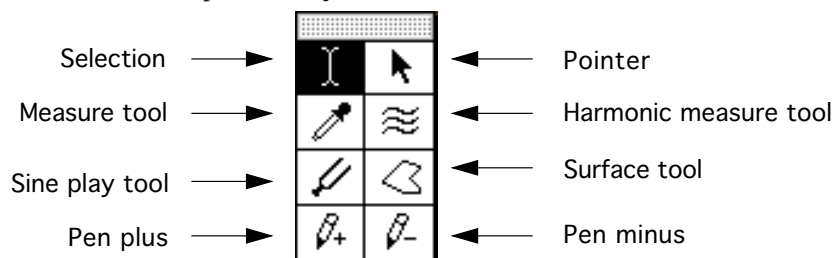
The three panes are separated by a thick grey line, the pane separator. Move the cursor above a pane separator: you'll see its shape change into two triangles. Now click with the mouse key in the pane separator, move it up or down and then release the mouse key: the pane will resize accordingly. This way you can resize each of the panes and even hide one of the panes when you move the pane separator against the window border.



The scroll bar at the bottom of the window lets you scroll the sonogram. Grab the handle in the scroll bar with the mouse and move the sonogram directly with the mouse in either direction. On a Power Macintosh the sonogram will follow in real-time; on slower computers refreshing the sonogram image might be slower.

The sonogram shows the frequency contents of the sound in time. The sound is decomposed into frequency components and the sonogram displays the evolution of the amplitudes of these components in time. The grey scale corresponds to the amplitude: white means a low amplitude, black means a high amplitude (when using the default white to black scale; the color scales are set with the color palette).

You should also find a toolbox palette on your screen. Here it is, with the names of the tools.



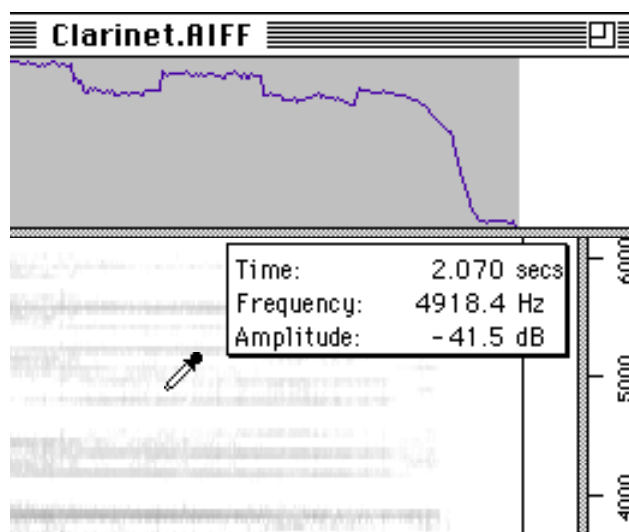
As an example, let's see what the Measure tool does.

Select the Measure tool by clicking its icon. Moving the cursor above the sonogram will change its shape into a pipette.

If you click in the sonogram an inspector window is shown on the screen. This window has three fields:

- *Time* tells you the time of the clicked point in the pane;
- *Frequency* tells you its frequency;
- *Amplitude* indicates the amplitude of the point in the sonogram (in decibel).

If you move the cursor vertically while holding down the mouse key, you'll see that the curve in the short-time spectrum pane is being traced. The displayed curve is actually one slice of the sonogram pane at a particular time. The amplitudes that are shown in shades of grey in the sonogram pane are shown here as an amplitude-frequency curve.

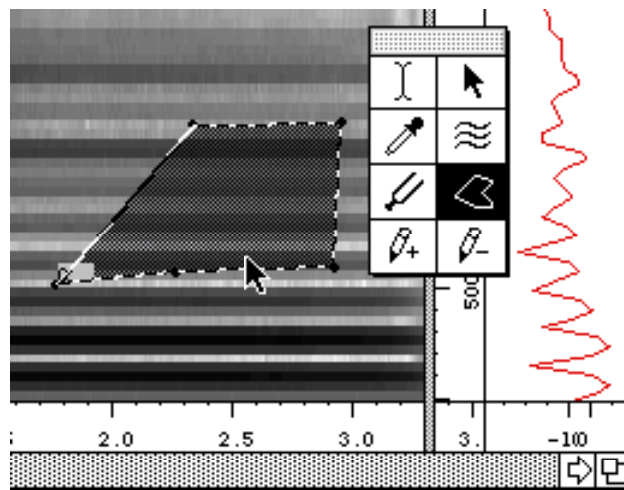


The other tools of the toolbox are discussed in Part 2 and 3 of this manual.

The sonogram filter

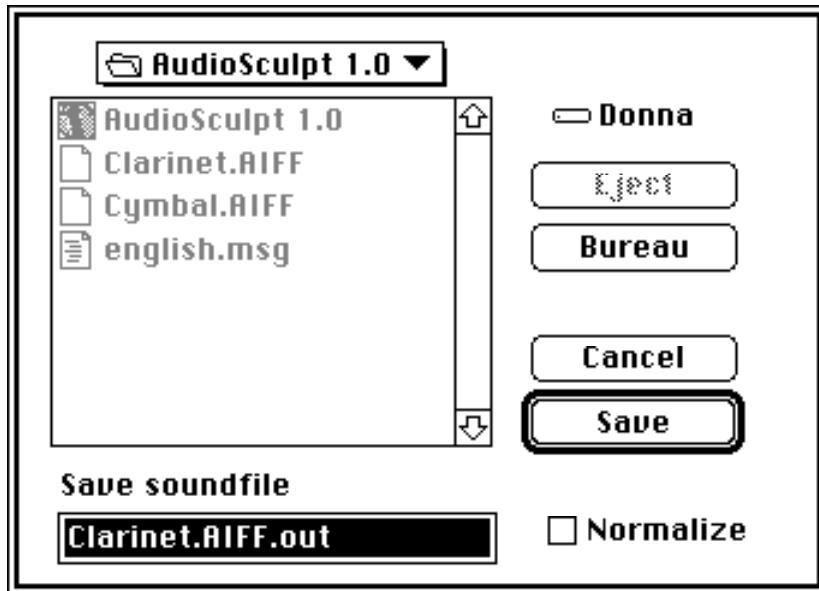
Take the surface tool and draw a line freely on the sonogram while holding down the mouse key. When you release the button, the line will be automatically closed. The surface that has been chosen is shown in grey. Here is how to draw a surface:

Here is another way to use the surface tool. Click once in the sonogram with the surface tool: one end of a line is fixed, the other end of the line can be moved with the mouse. Click a second time: the endpoint of the first line is fixed and a second line is drawn. Continuing this way you gradually draw a polygon. Double-clicking will close the surface.



Now that you can draw surfaces we'll continue with what it's designed for: filtering. Start selecting a surface with the pointer tool and type a number, 10 for example, on the keyboard. The number will be written in the label attached to the surface. The number in the label specifies a gain expressed in dB. All the frequencies that are covered by the surface will be amplified according to this gain when the sonogram filter is executed. You can type a negative gain also, type -10 for example, and the area that is covered will be attenuated.

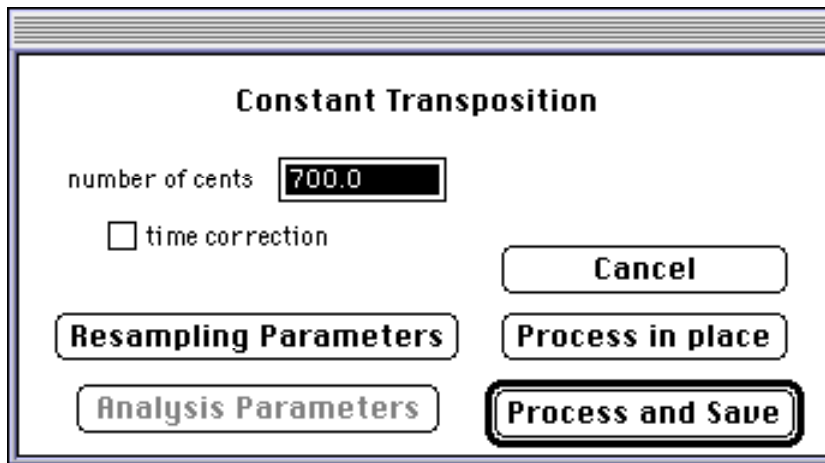
Give a gain to every surface you've drawn and then choose **Sonogram Filter** in the **Processing menu**. You'll be asked to give a name to the new file in the dialog box and then click the **Save** button. A process window comes on the screen and shows you the progress of the computation. The resulting sound will be displayed in a new window.



Just one last word before we finish this tour. You can save the sonogram by choosing **Save** in the **File** menu. The name of the new file will be asked in a dialog box. The analysis file will be saved separately from the soundfile. If you open the analysis file AudioSculpt will automatically open the soundfile with it. AudioSculpt is able to find back the soundfile even if you have renamed it or moved it into another folder. Through several examples, the next paragraphs will explain some features of the AudioSculpt interface.

‘Process in place’ or ‘Process and save’: Transposition

Open a soundfile and choose the **Transposition** command in the **Processing** menu. This command lets you shift the sound to a higher or a lower frequency. A dialog like the one in the figure below appears on the screen. The interval with which the sound has to be transposed can be written in the text zone besides 'number of cents'. 100 cents corresponds to the interval of a half tone. Type for example 700 cents, which corresponds to the interval of seven half steps, i.e. a perfect fifth. Now let's take a look at the options in the dialog box since most of the dialog boxes of the processing commands have them in common.



The name of the selected processing technique is marked at the top of the window. In this case it should say **Constant Transposition**. The right side of window shows three buttons one above the other: **Cancel**, **Process in place** and **Process and save**.

The **Cancel** button cancels the command. The dialog disappears and no processing takes place. The **Process in place** button starts the processing. The original soundfile will be overwritten with the resulting soundfile. This means the original soundfile will be lost unless you choose **Undo** in the **File** menu. The **Process and save** button equally start the processing but the resulting soundfile will be saved in a new document. The original soundfile will not be modified. The name and the format of the new document is asked in a new dialog. If the name of the new document already exists, AudioSculpt asks if this new document should replace the already existing one.

The right side of window shows two buttons: **Resampling Parameters** and **Analysis parameters**.

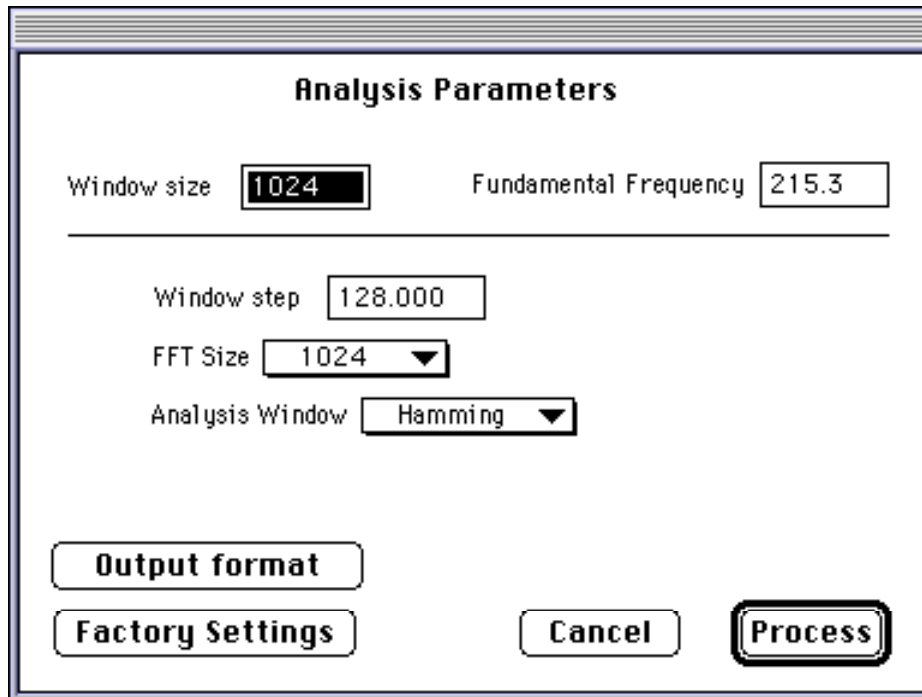
The resampling uses an interpolation function to generate the transposed sound. The length of this function can be set with the **Resampling Parameters** button. Further explanation can be found in Part 3.

Most of the processing and analysis command have an **Analysis parameters** button. In case of the transposition the button is only accessible if you mark the time correction option. The button brings up the **Analysis Parameters** dialog. In this dialog you can set the parameters for the FFT analysis that are used in the processing. A description of the parameters of the FFT analysis is given in Part 2.

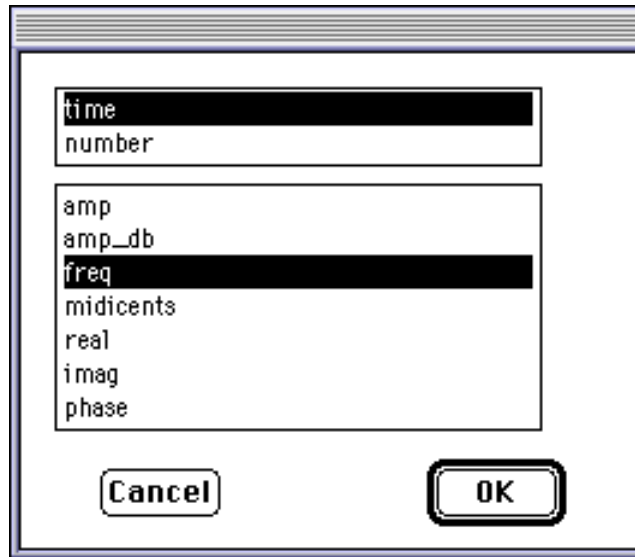
If the time correction option is marked, AudioSculpt will expand or compress the resulting soundfile, so it has the same length as the original sound.

The analysis modules

AudioSculpt offers several analysis modules based on algorithms that were previously developed at Ircam. These analysis modules do not change the soundfile but bring out certain characteristics of the sound. To give an example of such an analysis, we will execute a short FFT analysis. Select a small part of the soundfile, then choose the **FFT Analysis** command in the **Analysis** menu. You should see the following dialog.



The analysis here is the same as the FFT analysis of the sonogram except that the results of the analysis are saved in a text file. These text files can be very long and are meant to be read by other applications. You can define what data will be saved and in what format. Click the **Output Format** button and a window pops up.



It has two frames: the top one says **time** and **number**, the bottom one **amp**, **amp_db**, **freq**, **midicents**, **real**, **imag** and **phase**. You can select one of the fields by clicking on them. If you click a second time you unselect it. The selected fields will be exported to the text file. In this example we'll select **time**, **number**, **freq** (frequency) and **amp_db** (amplitude measured in dB). Click **OK**. Select the default settings for the analysis parameters and launch the analysis. A dialog box will pop up and let you choose a home folder and a name for the result file. During the processing, a window shows you the progress of the computation

Open the resulting file with AudioSculpt, if the file is less than 20000 characters, or with a text editor, if the file is longer than 20000 characters. The file should look like this:

0.550635	513
20.009872	0.000000
32.208248	122.701393
39.102013	173.265961
31.840477	213.133362
13.864026	8.929436
31.077595	132.973312
40.107166	175.546097
34.137051	217.319733
17.722542	349.320221
21.023300	495.663269
32.389496	522.341614
26.851715	559.427124
16.214020	356.371521
11.621792	554.405151
25.198784	524.885559
20.740185	566.830200
9.803786	696.411743
8.409177	842.804688

The first line shows the time of the analysis (offset in the soundfile measured in seconds) and the number of analysis points (in this case the number of points is equal to the FFT size divided by 2). As of the second line, the result of the analysis is given in the columns. The first column shows the amplitudes of the spectrum measured in decibels and the second column shows the instantaneous frequencies of the spectral components.

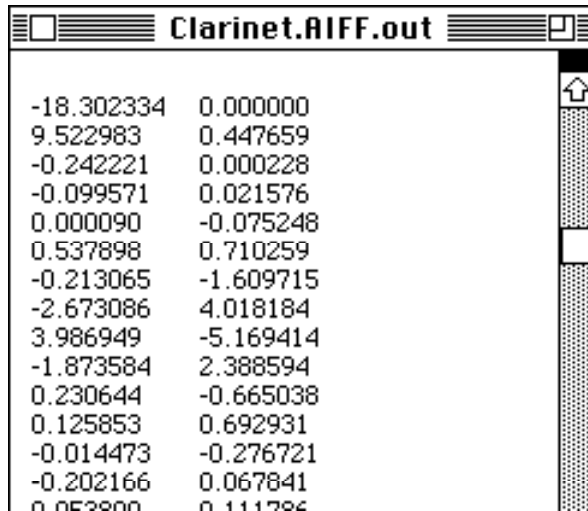
The command line

Most of the AudioSculpt commands can be launched using such a command line. A command line is a text line which closely resembles an UNIX-like instruction. In appendices A and B you find a list of all commands and their options. The folders in which the parameter and console files are looked for or saved can be set with the **Set Default Folders** command.

Now, choose **Use Command Line** in the **Processing** menu. A window will be displayed. Type the following text in the text zone of the command line window:

```
svp -v -t -SClarinet.AIFF -B0.0 -E0.5 -Afft -U -Oa -Whanning -N4096 -M4000 -I1024 Clarinet.AIFF.out
```

Click **Process**. The console window is brought onto the screen showing the sound format and the parameters of the analysis. When the analysis is completed the text file should look like this:



-18.302334	0.000000
9.522983	0.447659
-0.242221	0.000228
-0.099571	0.021576
0.000090	-0.075248
0.537898	0.710259
-0.213065	-1.609715
-2.673086	4.018184
3.986949	-5.169414
-1.873584	2.388594
0.230644	-0.665038
0.125853	0.692931
-0.014473	-0.276721
-0.202166	0.067841
0.052000	0.111796

You'll find the values of the complex spectrum in columns (see Part 2 to know more about complex spectrum).

You can copy and save a command line in a text file that you can create from AudioSculpt's command **File>New**, and thus create a library of command lines.

Processing and parameter files: formant filter

Open a soundfile and then open a new text document in the **File** menu. In the **Processing** menu you'll find four filter types:

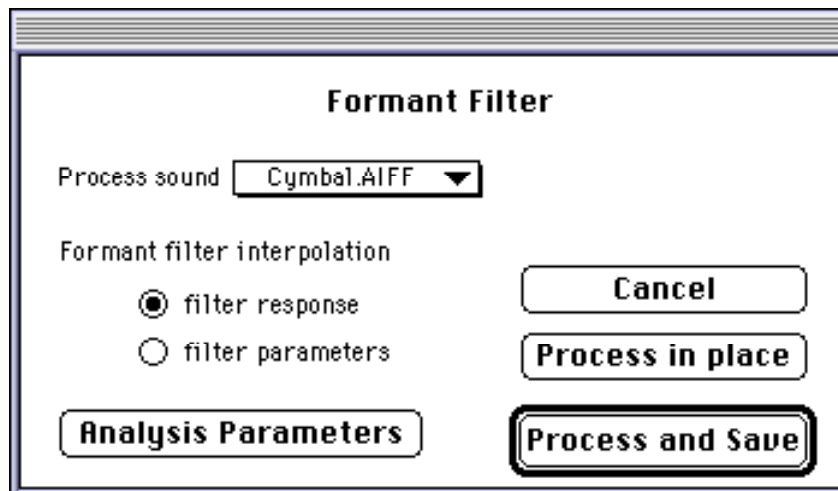
- breakpoint filter,
- band filter,
- surface filter,
- formant filter.

Caution: these items are activated (shown in black) **only** if a text file is active.

Each of the four filters needs specific parameters which are specified from within a text file. In this paragraph we will give an example of the formant filter.

formant filter parameters									
0.0	1	440	0	50					
1.0	2	440	0	50	880	-2	100		
2.0	3	440	0	50	880	-2	100	1760	-2 200

Copy the values given in the figure above into your text window. On each line, these values should be separated by at least one space or Tab and each line should be terminated by a carriage return (use the Return key). When you've copied the values choose the **Formant Filter** option. In the dialog box you specify to which of the open soundfiles you wish to apply the filtering (choose Cymbal.AIFF). Be sure that the option **filter response** is marked and then click **Process and Save**. Give the new soundfile a name and press **OK**. The new sound will appear on the screen in its own window.



A formant filter is a filter that is made up of second order filters. Each second order filter is defined by specifying its central frequency, its bandwidth and its gain. These parameters can vary in time. The first column of the parameter file says when to apply the filter parameters (measured in seconds) that are written on the same line. The second column says how many second order filters are used. On the same line you then find the parameters for each second order filter: the central frequency, the gain and the bandwidth, in that order and for each filter. To move from one parameter value to the parameter value on the next line, a linear interpolation is done. So between 0 seconds and 1.0 seconds a linear interpolation is done between a filter with 1 formant at 440 Hz and a filter with 2 formants at 440 Hz and 880 Hz.

You can use parameter files for the filters mentioned above, but also for time-varying transposition, for time-varying time scale modification and for time-varying general cross synthesis. The parameters that you need to define are discussed in Part 3.

What's next?

Part 2 of the manual discusses the analysis modules. In Part 3 the processing functions of Audio-Sculpt are detailed. And in Part 4 the menus are displayed and a brief description of the commands is given.

Part 2

Analysis

1 Analysis

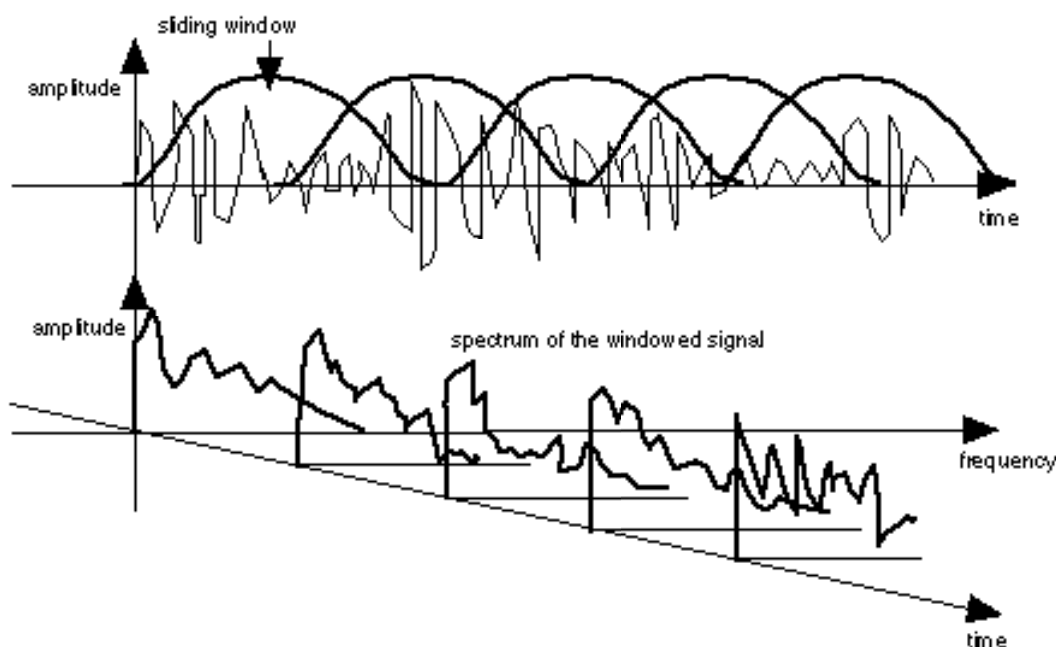
General principles of the short-time spectral analysis

Short-time spectral analysis is a class of several sound analysis techniques that have some features in common. The analysis consists of dividing the sound into small segments and looking at the frequency contents of each segment by creating a spectral representation of it. This spectral representation can be obtained using one of several spectral analysis techniques. Each segment is obtained by multiplying the signal by a window in the time domain. To obtain the next segment, the windowing operation is repeated after having shifted the window a certain distance along the time axis. Each of the segments obtained in this way is processed by an algorithm which calculates its short-term spectral representation.

One of the short-time spectral representations is the Fourier transform which gives the amplitude and phase

of the signal as a function of frequency, that is, the energy distribution of the signal frequency, like the sonogram.

The short-term spectral representation is a useful representation for two reasons: on the one hand, it is an intuitive representation for the user, because it functions much in the same way as the human auditory system, which detects the frequencies that make up a sound at any given instant. On the other hand, because of the decomposition of sound along two variables, time and frequency, it separates the temporal and frequential information of the sound. This allows the temporal information to be manipulated independently of the frequential information. Because in many cases the goal of such an intermediate representation is to be able to modify the signal's properties, it is important to know to what degree the short-time spectrum gives an accurate description of the signal over the chosen portion.



Window size

The number of samples in the windowed segment is named the window size. This windowing operation is performed by multiplying the signal by a fixed function of time at a certain time index within the sound. The role of this operation is to pick out a small segment of the signal, while at the same time it guarantees sufficient spectral and temporal resolution.

The short-time spectrum gives an accurate description of the characteristics of the signal as long as the sig-

nal is stationary over the length of the segment delimited by the window. In general, this is the case for short durations during which the characteristics of the amplitude-time curve for the sound (amplitude envelope, length of the period, shape of the wave form, etc.) do not change appreciably.

If the segment has a stable frequency content, the frequency information of the sound segment is contained in the short-time spectrum, and the temporal information is reflected by the changes in the frequency composition over successive analyses.

If the short-time spectrum is calculated over a portion of signal that is non-stationary, the changes in time will not be accurately registered in the successive analyses. If the window size is large, the short-time spectrum corresponds to a large section in the soundfile. If the amplitudes of the frequency components change over this section, the change will not be reflected in the short-time spectrum. If the sound evolves rapidly, a short window size is needed to accurately register the changes. So, the window size conditions the temporal resolution of the analysis.

The frequency resolution also depends on the window size. If a detailed spectral analysis is desired a large sound segment should be used, and thus a large window. If the window is small, it is hard to decide precisely what the frequency components of the sound are. This is reflected in a less precise spectral analysis. Therefore, a compromise should be made.

Window step

The next window, or analysis frame, is obtained by shifting the window (in time) by an amount called the window step. In general, and for reasons that come from the theory of spectral analysis, successive windows should overlap; that is, the window step is always smaller than the window size and should not be bigger than $1/4$ of the window size. In AudioSculpt, the window step is equal to $1/8$ of the window size, by default.

The temporal resolution of spectral evolution also depends on the window step. In other words, the details of the temporal evolution of a sound are best preserved when the window step is small.

It is possible to control the window size and window step separately, so as to adapt the window size to the frequency composition of the sound, while at the same time having control over the temporal evolution. For example, it is possible to use a large window size for a low frequency sound (a low frequency has a long period, so the window should be long), but set a small window step in order to resolve transient components.

Choice of window

There are different kinds of windows. The choice of a particular type is determined by its effect on the spectrum of the signal. The obtained short-time spectrum is not the exact spectrum of the signal, but the convolution of the spectrum of the signal with the window's spectrum.

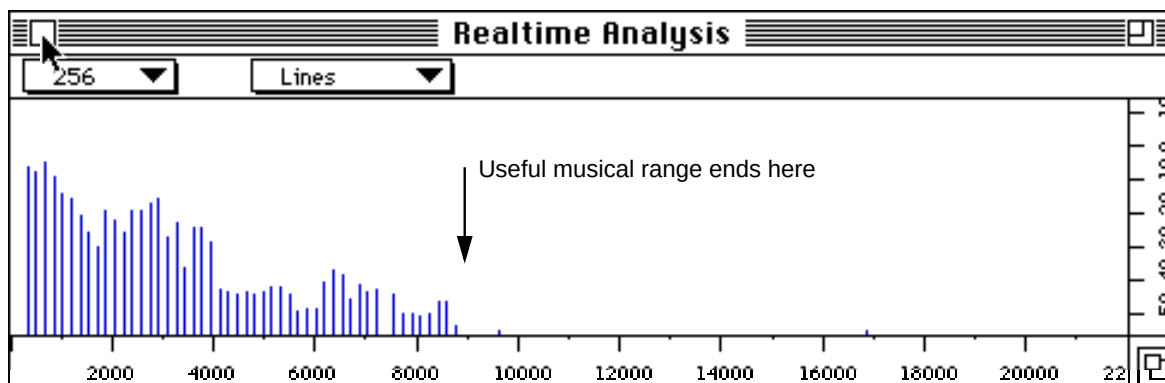
Spectral analysis

There are different short-time spectral analyses that allow the frequency contents of a sound to be evaluated, including the Fourier transform, linear predictive analysis, discrete cepstrum, etc. In this version of AudioSculpt, you'll find the Fast Fourier transform and the linear prediction analysis. Future versions of AudioSculpt will include the discrete cepstrum analysis. The parameters used in the analysis modules are, among others, the type of spectral analysis (FFT, LPC), the size of the FFT, the window size and the window step.

In the next paragraphs we will give a brief description of the FFT and LPC analysis. This description is merely an introduction to explain the basic features of these analysis techniques and their parameters. For a more detailed description, the reader is referred to a theoretical work on signal processing.

Fast Fourier transform (FFT)

The FFT is the algorithm used to compute the Fourier transform, which converts a windowed signal into a spectrum. The FFT gives, over a certain time interval, the decomposition of the sound into a sum of narrow-band signals, centered around successive frequencies that are linearly spaced over the frequency range going from 0 Hz to $SR/2$ (SR stands for the sampling rate of the soundfile). The number of frequency components in which the sound is decomposed is equal to half the FFT size.



Amplitude and phase spectrum

The FFT computes a complex value for each frequency component. These complex values can be represented as a real and imaginary part or can be transformed into an amplitude and a phase value. The amplitude corresponds to the energy of the spectral component. The phase deals with the instantaneous frequency, that is, the small deviation of the frequency of the component relative to its central frequency. The series of complex values of every frequency component is called the complex spectrum and to every complex spectrum corresponds one amplitude and one phase spectrum.

The short-time Fourier representation makes it possible to follow the time evolution of the amplitudes and phases of the frequency components of a sound. Most filters are designed to manipulate the amplitude spectrum. The only exceptions are the breakpoint filter and the generalized cross synthesis, as they actually manipulate the phase spectrum. Changing the way the phase spectrum evolves in time will result in a modification of the frequencies of the components.

Windowing effects

Windowing has two kinds of side effects on the FFT spectrum. These side effects depend on the window size and on the type of window. The first side effect is in the frequency resolution and the second side effect is that windowing can introduce spectral components where there weren't any previously.

Because of the windowing, the sound spectrum is convolved with the spectral representation of the window signal. This has an effect on the spectral resolution. Two spectral components that have only a small difference in frequency between them can be merged into one spectral peak in the short-time spectrum and it becomes impossible to detect them separately. The resolution depends on both the window size and the window type. The spectral components of a sound are best isolated when the window size is large. The number of points in the window should be greater than or equal to a multiple of the number of points in a period.

This multiple depends on the type of window. For a Hamming and Hanning window, it is equal to 4. The multiple for the Blackman window is bigger. As a general rule, in order to obtain an acceptable analysis for pseudo-periodic sounds, it is recommended that a window size of perhaps 4 or 5 times the period of the sound be used. AudioSculpt takes a window length of 5 times the period length. If the sound is not periodic, the window size is determined by the distance between the closest frequency components that one wishes to separate.

For example, if the closest spectral components have a difference between them of 100 Hz and the sound is sampled at 44100 samples per second then the window size should be at about 2205 samples long. Indeed, 100 Hz corresponds to a period of 10 msec, which on its turn corresponds to 441 samples. Multiplying by five gives the final result.

Your goal in using AudioSculpt will usually be the modification of the sound's spectrum. Most of the processing techniques offered by the program involve the conversion of the sound into spectrum, a procedure that is transparent for the user. In this process, however, it is desirable to have the least possible spectral distortion. When the window is enlarged, the window effects are diminished, and the spectral components are better resolved. The window size thus has a direct influence on the quality of the spectral data obtained. On the other hand, the spectrum must not be so large that the resolution of the time evolution of the spectrum is compromised. The bigger the window size, the longer is the sound interval that is analyzed. In other words, the temporary evolution within that sound interval are averaged into one short-time spectrum. It makes it hard to decide at what time exactly an event took place. With smaller windows on the other hand the events are better captured.

Because of windowing, the spectrum of the sound is convolved with the spectrum of the window and peaks can be introduced or reinforced in the sound spectrum. This effect mainly depends on the type of window used. The effect is smaller when the Blackman window is used than when the Hanning or Hamming window is used.

What window size and window type is best for a given sound is a delicate matter. If the sound has spectral components that are close in frequency, it is better to choose the Hanning or Hamming window to guarantee a better spectral resolution than with the Blackman window. On the other hand, if the sound has only a few, distinct spectral components, it is better to use the Blackman window to avoid as much as possible the introduction of spurious peaks. This is but a mere indication. If the intermediate spectral representation is manipulated in view of a synthesis, the choice of the window type and the window size can also depend on the type of processing applied.

FFT size

When using the FFT algorithm, the number of points in the windowed segment must be a power of two. But

for a pseudo-periodic sound, the window size is not necessarily equal to a power of two. If the window is enlarged to a power of two, the resolution in time of the short-time spectrum is diminished. To resolve this problem, a window size is initially chosen (5 times the length of the period, for example), and then is padded with enough samples with value zero so that the next-highest power of two is attained.

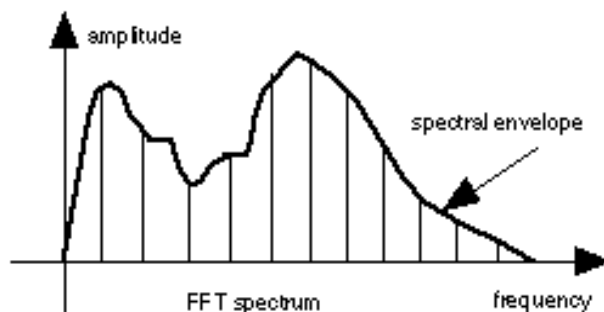
Adding zeros to a windowed signal has no effect on frequency resolution (separation of the sound's frequency components). The spectral resolution, however, is improved, because the spectrum has been more finely sampled. Increasing the FFT size makes the spectrum easier to inspect.

The size of the FFT determines the spectral resolution, which is the distance between two frequency bins in the spectrum. Spectral resolution is given by the ratio of sampling rate to FFT size.

Linear Prediction (LPC)

The FFT and LPC are not equivalent because they compute a spectral representation that does not contain the same information. The FFT contains all the spectral information, while in the LPC spectrum, only the spectral envelope is retained. The implementation of LPC is based on a program by Ph. Depalle. The spectral envelope concerns the contour of the amplitude spectrum and not so much the spectral components. The knowledge of the spectral envelope allows the determination of the maxima of the envelope which are called the formants. The position of these formants plays, for example, an important role in the recognition of vowels in speech. The spectral envelope can also be used to filter another sound. We will come back to this option in the section on cross synthesis source-filter in the Processing chapter.

The LPC technique calculates a filter that has a frequency response which closely resembles the spectral envelope of the sound. The transfer function that represents the filter is defined by a set of poles, which go in pairs. Each pair of conjugate poles represents a formant in the spectral envelope. The number of poles (equal to twice the number of formants), controls the smoothness of the curve. The number of poles in the LPC analysis can be specified by the user, so as to increase the precision of the spectral envelope. Thirty poles are used by default. The higher the order of the analysis, the more detailed the envelope is described and the more individual spectral components are modelled. A detailed description is not always necessary, one might be interested in the global structure of the envelope and only notice the most significant formants (in case of a cross synthesis for example). Once the poles of the filter are known, the filter response (or spectral envelope) can be calculated using the FFT algorithm. So, you can specify the number of points in which the spectral envelope is shown by setting the FFT size. Here is a figure of the FFT spectrum and its spectral envelope.



Analysis display: sonogram or text file

AudioSculpt offers two ways of inspecting the results of the spectral analyses. The results can be displayed as an image on the screen in what is called the sonogram (see also the Introduction: A quick tour). The results can also be displayed in ASCII format and saved in a text file.

The sonogram

In the Analysis menu you find the command **Sonogram** analysis. It brings a dialog on the screen:

Analysis Parameters

Window size Fundamental Frequency

Window step

FFT Size

Analysis Window

Synthesis Window

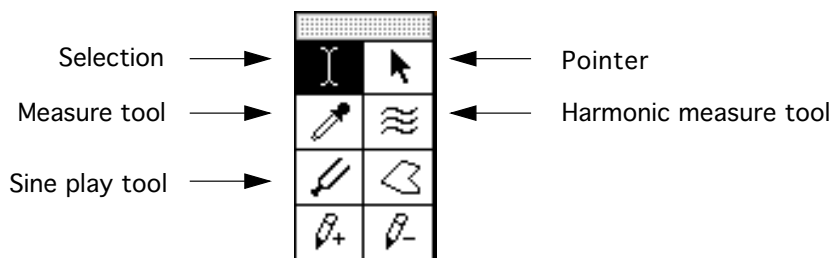
The dialog lets you specify the analysis parameters and the type of spectral analysis. Here is a brief survey of the parameters:

Window size:	the window sizes determines the number of sound samples in each analysis.
Fundamental frequency:	allows you to compute the window size automatically. If you specify the window size, the fundamental frequency that is calculated will be displayed. The window size is set to five times the length of the period that corresponds to the fundamental frequency.
Window step:	the window step determines the time interval between two successive analyses. The time interval is measured in number of samples.
FFT Size:	the FFT size determines the number of points of the analysis. The number of points in which the spectrum is displayed is always one more than the FFT size divided by 2.
Analysis window:	in the analysis window menu you can specify the window type used in the analysis. Three types of window are proposed: Hanning, Hamming and Blackman.

If you use the LPC analysis the number of poles have to be set. The number of poles should be between 10 and 200; by default 30 poles are used.

The **Factory Settings** button sets the parameters to the default value. The sonogram analysis concerns the whole soundfile, even if you have selected but a small part. On launching the analysis, two new panes are added to the window beneath the existing sound pane: the left pane is the sonogram pane, the right pane is the short-time spectrum pane. You can stop the analysis with the **Stop** button in the progress window or by selecting **Sonogram Analysis** in the **Analysis** menu.

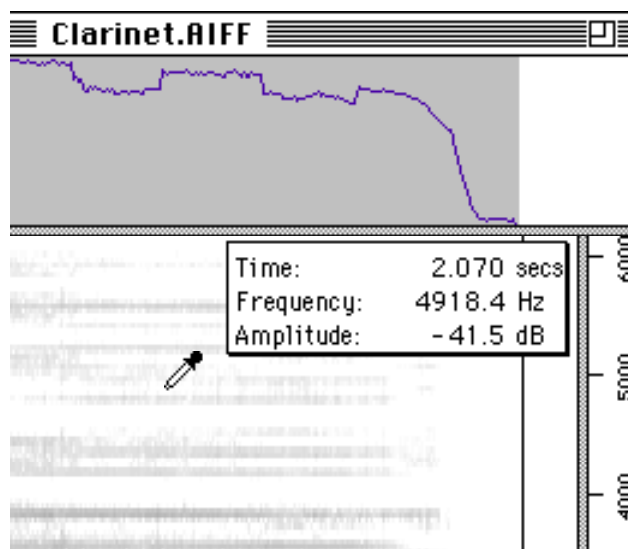
The toolbox



After the analysis is done a toolbox is brought on the screen. Some of the tools help you inspecting the sonogram; others can be used to manipulate the sonogram to create a new sound and will be discussed in the Processing chapter. Here follows a short description of the analysis tools:

The **Measure tool** gives you the coordinates of the cursor position in the sonogram. Clicking with the

mouse in the sonogram shows you the exact time, frequency and amplitude of the clicked point. The short-time spectrum pane displays the amplitude frequency curve at the selected time (the displayed short-time spectrum is a scan of the sonogram along the frequency axis). When you activate this tool, the inspector window is displayed:



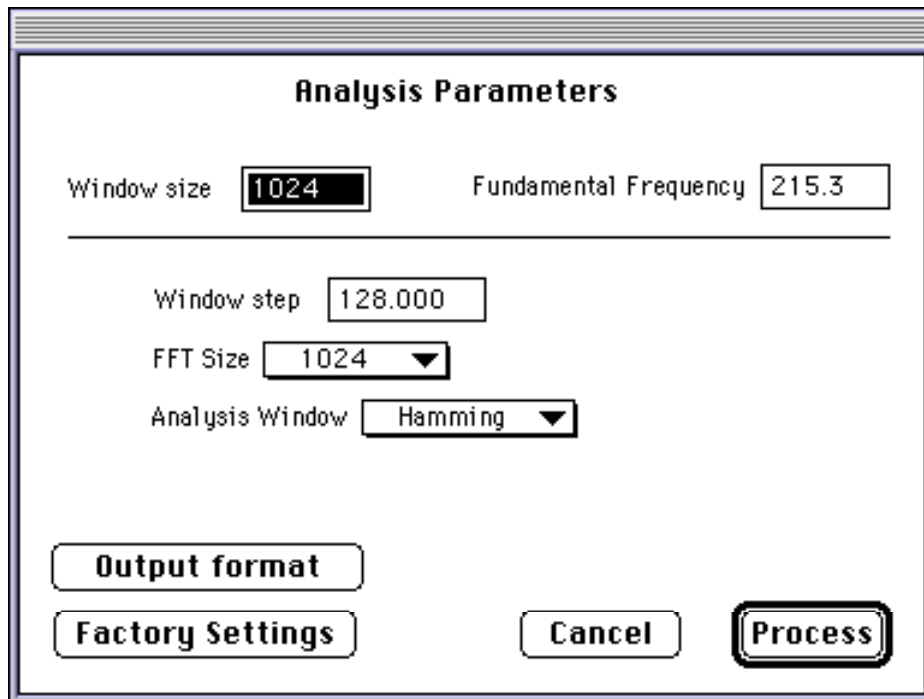
The **Harmonic measure** tool shows the first 18 harmonic partials of a fundamental frequency in the form of horizontal lines on the screen. The position of the cursor defines the fundamental frequency. This tool helps you scan for the exact position of the partials in the sonogram by comparing them with the harmonic spectrum of the partial measure tool. The frequency scale on the right of the sonogram can help you in measuring the frequencies.

The **Sine play tool** plays a pure sine tone with a frequency and amplitude corresponding the clicked point in the sonogram. The amplitude depends on the grey scale of the clicked point. You can adjust the volume with the color palette.

Caution: this tool works only on Power Macintosh.

Text file output

The results of the FFT and LPC analyses can also be saved in a text file. These text files are mainly meant to be read by other applications like PatchWork, Max or Excel since these files can be quite voluminous. The analysis can be applied to a small region of the sound. To do so, select the region before selecting **FFT Analysis** or **LPC Analysis** in the **Analysis** menu. For the FFT a dialog will be displayed.



Analysis Parameters

Window size Fundamental Frequency

Window step

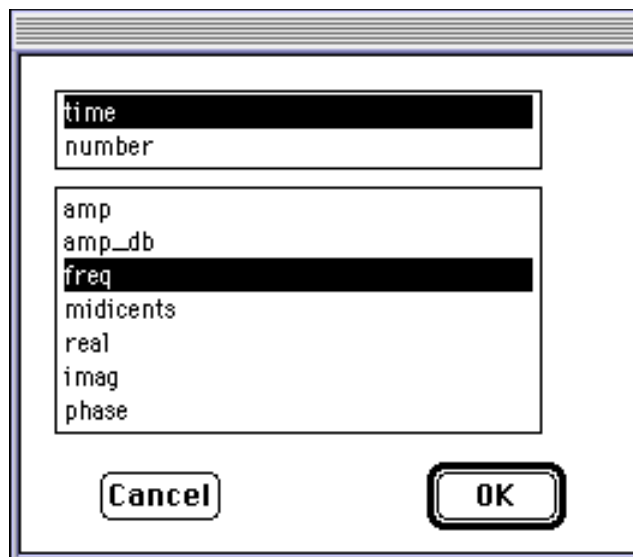
FFT Size ▼

Analysis Window ▼

Output format

Factory Settings **Cancel** **Process**

It lets you specify the parameters. With the **Output format** button you can bring up the dialog displayed below. Here you can specify the format in which results will be written to the text file. It has two frames: the top one says **time** and **number**, the bottom one **amp**, **amp_db**, **freq**, **midicents**, **real**, **imag** and **phase**.



time
number

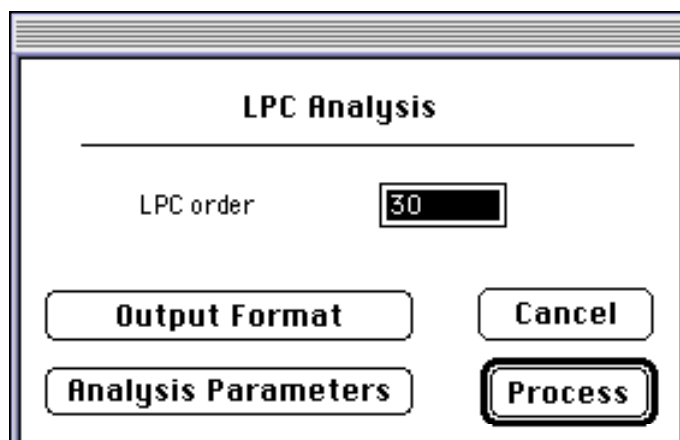
amp
amp_db
freq
midicents
real
imag
phase

Cancel **OK**

You can select one of the fields by clicking on them; a selected field is shown highlighted. If you click a second time you unselect it. The selected fields will be exported. If time is selected, the resulting file will

show the time of the analysis (offset in the soundfile measured in seconds) in the first line of the results. If number is selected, the first line will show the number of point of the analysis. The number of points is one more than the FFT size divided by 2. The result of the analysis are given in columns. If **amp** is selected, one column will show the amplitudes of the spectrum envelope. If **amp_db** is selected, one of the columns will show the amplitudes measured in decibels. With the **freq** field selected, the frequencies will be written in a column and the midicents field gives these frequencies in midicents. In case of the FFT, the instantaneous frequencies are exported. That is, the deviation of spectral components from the central frequencies are taken in account. For the LPC analysis the frequencies of the analysis points (the central frequencies) are exported. The parameters **real** and **imag** give, respectively, the real and the imaginary component of the complex spectrum. If **phase** is selected the phase spectrum will be exported.

The dialog for the LPC analysis looks like this.



You can set the order of the LPC, and with the **Analysis** button the parameter values for the FFT algorithm can be given as well. The **Output Format** shows most of the options that were discussed for the FFT analysis.

Detection of partials: peak detection and fundamental estimation

AudioSculpt offers two analysis modules that inspect the structure of the amplitude spectrum. The first one looks for peaks in the spectrum, the second one concludes on the fundamental frequency based on the peaks in the spectrum.

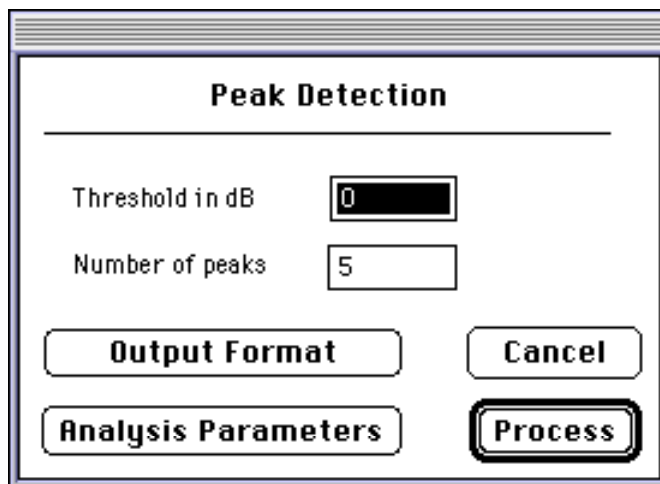
Both analyses save the results in a text file if the sound has no sonogram open; if a sonogram is shown, the results will be presented as a breakpoint curve on the screen.

Peak detection

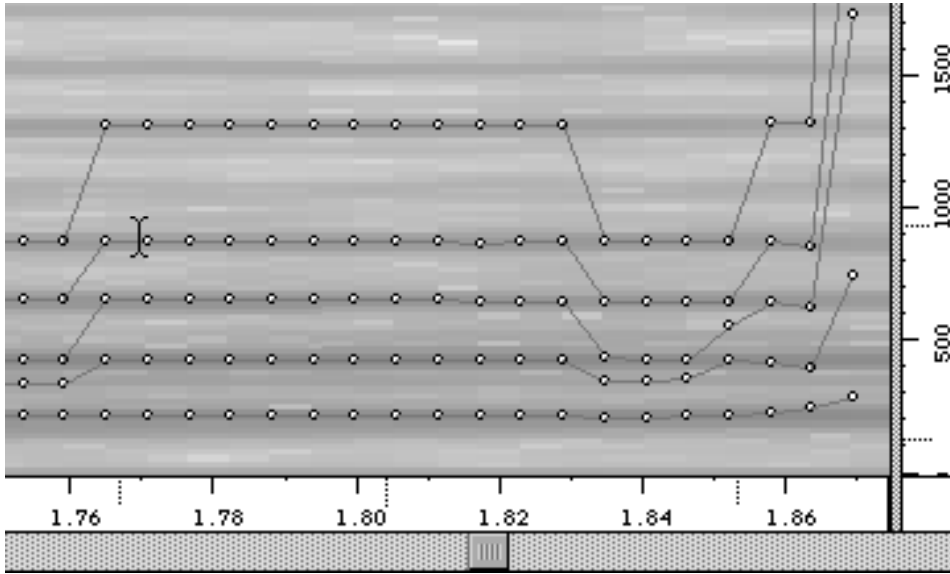
The **Peak Detection** command looks for the maxima in the spectra. It is based on research by G. Garcia, X. Rodet and Ph. Depalle. The algorithm looks for the most significant ones, and ignores the other, which are called "spurious peaks".

In each short-time amplitude spectrum the amplitudes of the maxima are detected. The differences in amplitude between the maximum and the neighbor points must exceed a pre-defined threshold. All peaks that do not satisfy this constraint are eliminated. Then you can ask to export only the peaks with the highest amplitudes. The number of peaks that will be exported can be defined before the analysis.

You can apply the analysis on a part of the sound by selecting it before launching the **Peak Detection** command in the **Processing** menu. The dialog for the peak detection is shown below. It lets you set the threshold and the number of peaks to be registered.



With the **Analysis Parameters** button you can set the values of the analysis parameters. If the window contains a sonogram, the results will be represented:



It shows the results as breakpoint functions on the sonogram: the results in each analysis window is shown as a small breakpoint on the sonogram and the breakpoints of successive analysis are connected with a line. With the pointer tool in the toolbox, you can grab a breakpoint and move it. This way, erroneous results can be modified by the user.

In case the window only has a soundfile open and no sonogram analysis has been performed yet, the results will be saved in a text file. Again, you can choose what results will be exported. Therefore, you click the **Output format** button, and then in the dialog you can mark the desired data. The formats of the results that can be chosen are:

- **time** (in seconds),
- **number** (the number of peaks desired),
- **amp** (the amplitude of the peak),
- **amp_db** (the amplitude of the peak in decibel),
- **freq** (frequency of the peak),
- **midicents** (frequency of the peak in midicents).

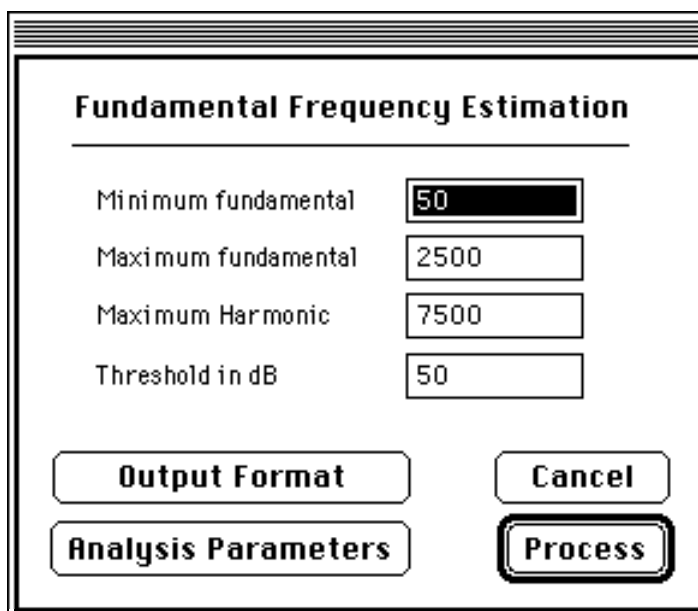
The peaks are ordered according to decreasing amplitude.

Fundamental estimation

The **Fundamental Estimation** command looks for a fundamental frequency in every short-time spectrum. The algorithm was conceived by B. Doval and X. Rodet. In every analysis window the peaks in the ampli-

tude spectrum are detected, then an estimation of the fundamental frequency is made based on these spectral peaks. You can set a threshold to eliminate spurious peaks: if the difference in amplitude between the highest peak and another peak in the amplitude spectrum exceeds the threshold, the peak is eliminated from the analysis. This reduction can be seen as a noise reduction technique. To decrease the calculation time of the algorithm and to help it in its search, you can specify a minimum and maximum fundamental frequency. The algorithm will look only for a fundamental frequency within the specified frequency region. Also you can set a maximum harmonic frequency. Spectral peaks with a frequency higher than this maximum frequency will not be taken in account.

The analysis can be applied to a selected part of the sound. Mark a region in the sound pane before choosing the **Fundamental Estimation** command in the **Analysis** menu. In the dialog window the analysis parameters can be set (**Minimum fundamental**, **Maximum fundamental**, **Maximum harmonic** and **Threshold in dB**).

A screenshot of a software dialog box titled "Fundamental Frequency Estimation". The dialog has a title bar at the top. Inside, there are four labeled input fields: "Minimum fundamental" with a value of 50, "Maximum fundamental" with a value of 2500, "Maximum Harmonic" with a value of 7500, and "Threshold in dB" with a value of 50. At the bottom of the dialog, there are four buttons: "Output Format", "Cancel", "Analysis Parameters", and "Process". The "Process" button is highlighted with a double border.

The default values are 50 Hz and 2500 Hz for the minimum and maximum fundamental, the default value for the maximum harmonic is 7500 Hz. The ratio Maximum harmonic/Minimum fundamental is limited for reasons of memory storage and calculation time. The parameters for the spectral analysis can be set using the **Analysis Parameters** button.

If the sonogram is displayed, the results will be shown on it as a breakpoint function. The breakpoints can be edited with the Pointer tool. If the sonogram is not displayed, the results are registered in a text file. You can then set the formats of the result using the **Output Format** button: **time** (in seconds), **number** (the number of peaks will be 1), **amp** (the amplitude of the peak fundamental frequency), **amp_db** (the amplitude in

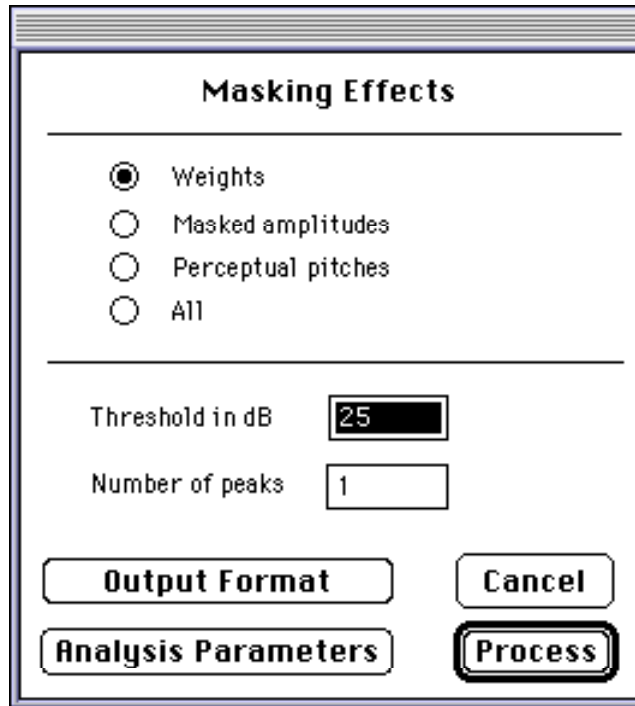
decibel), **freq** (the fundamental frequency), and **midicents** (the fundamental frequency in midicents).

Analysis based on perceptual criteria: masking effects

The **Masking Effects** command reduces the amplitude spectrum to a small number of frequencies which contribute to the tonal perception. The algorithm is based on work by Terhardt. It models the masking of spectral components which occur in our perception. Both the amplitude and frequency of the components are taken into account in measuring the masking effects. A spectral component can be masked and thus become perceptually less significant, because of the presence of a spectral component in the same frequency region. As the human ear is not equally sensitive to all frequencies, the perceptual importance of a spectral component also depends on its frequency.

First, the maxima in the amplitude spectrum are detected, using the same algorithm as in **Peak Detection** (see above). The algorithm then works on the remaining maxima. Processing calculates the amplitudes of the peaks after modelling the masking effects (**Masked Amplitudes**). The effects of masking are subtracted from a peak's original amplitude in order to obtain the masked amplitude. According to Terhardt's algorithm, a significant component has a positive masked amplitude (in dB). A component with a masked amplitude smaller than zero is considered as perceptually insignificant. Due to the interaction of several simultaneous spectral components, the frequencies that are perceived are not necessarily the same as the frequencies that were detected. These small frequency shifts of the spectral components are calculated in order to obtain the **Perceptual Pitches**. The third result, **Weights**, accounts not only for the masking effects but also for the spectral importance of the components. A weight bigger than zero is given to significant components; a zero weight is given to the insignificant components.

Choosing the **Masking Effects** command in the **Analysis** menu will bring up a dialog:



If a sonogram analysis has been performed, the spectral component with the highest weight is traced on the sonogram.

If no sonogram is displayed, the results will be registered in a text file. The dialog lets you determine what analysis will be performed: **Weights**, **Masked amplitudes**, **Perceptual pitches** or **All** of the preceding. The threshold for the peak detection can be set and you can specify how many peaks you want to be saved. Only the peaks with the highest amplitude will be given as a result. By default the highest peak is given.

The formats of the results can be set using the **Output Format** button: **time** (in seconds), **number** (the number of peaks desired). The next four describe the peaks in the amplitude spectrum: **amp** (the amplitude of the spectral peaks), **amp_db** (the amplitude of the peaks in decibel), **freq** (the frequency of the peaks), **midicents** (the frequency of the peaks in midicents). According to the requested analyses, the weights, the masked amplitudes, the perceptual pitches or all of these can be saved. The parameters for the spectral analysis can be set using the **Analysis Parameters** button.

Part 3

Processing

1 Processing: overview

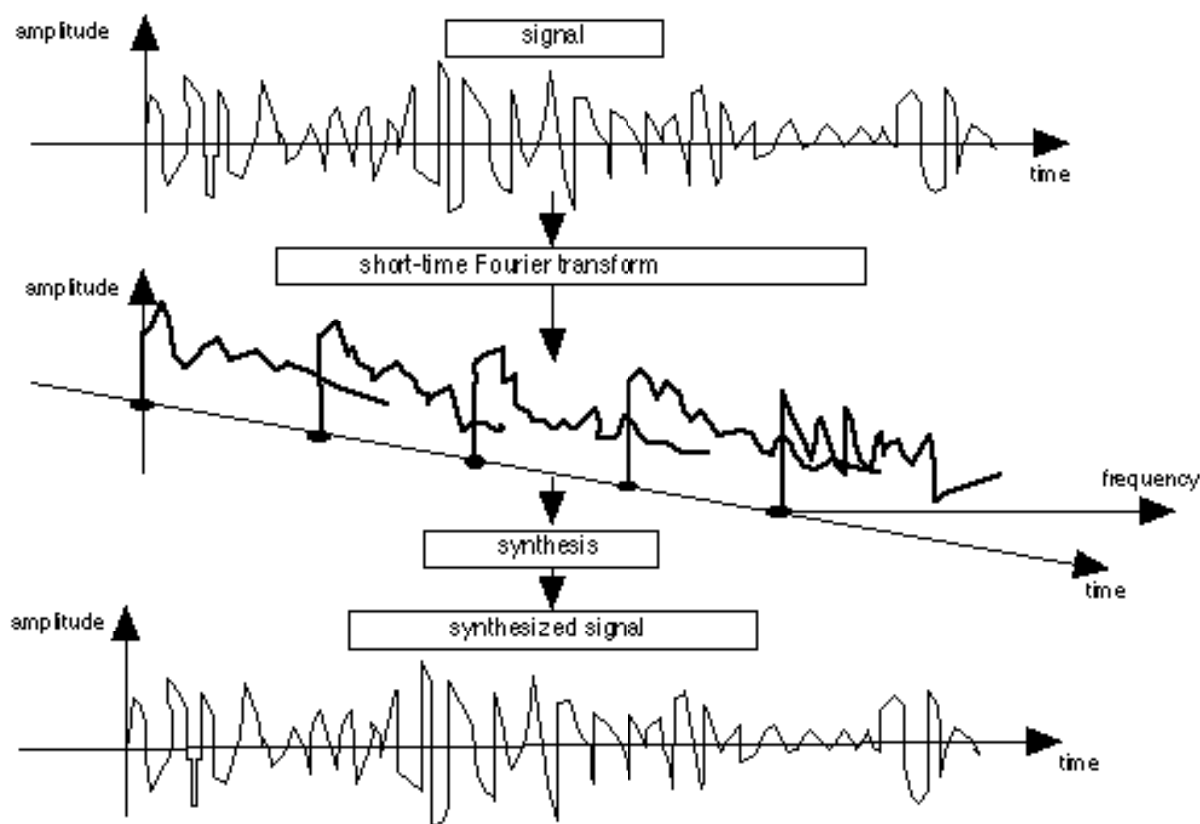
Most of the sound processing modules use an intermediate short-time spectral representation of the sound. This means that the first step of the processing consists of an analysis that generates this intermediate spectral representation. The second step consists of modifying this representation and finally the resulting soundfile is reconstructed using the inverse transformation of the analysis step (re-synthesis). In order to perform the spectral analysis, the sound signal is “windowed”. This windowing operation is performed by multiplying the signal by a fixed function of time at a certain index within the sound. The role of this oper-

ation is to pick out a small segment of the signal. The series of samples in the windowed segment is transformed into a spectrum, by applying a Fourier transform. As in the figure below, for each window of signal, there is a corresponding spectrum (see Part 2, Analysis). The short-time spectrum is a spectral representation of the segment of signal selected by the window and gives, over a certain time interval, the decomposition of the sound into a sum of sinusoids, each of which is characterized by its amplitude and phase.

The fundamental characteristic of the Fourier transform is that it gives an intermediate representation of a signal in which frequential and temporal information are separated. The intermediate representation, in the form of a series of Fourier transforms, makes apparent the distribution of the energy of the signal by frequency, as well as the way in which this distribution changes with time. Before the signal is reconstructed, the user can work directly on this intermediate representation so as to modify the original signal.

Since this representation contains two types of information, temporal information and frequential information, two different kinds of modifications must be distinguished, temporal modifications and frequential modifications.

Here is an illustration of the analysis/synthesis procedure for the phase vocoder:



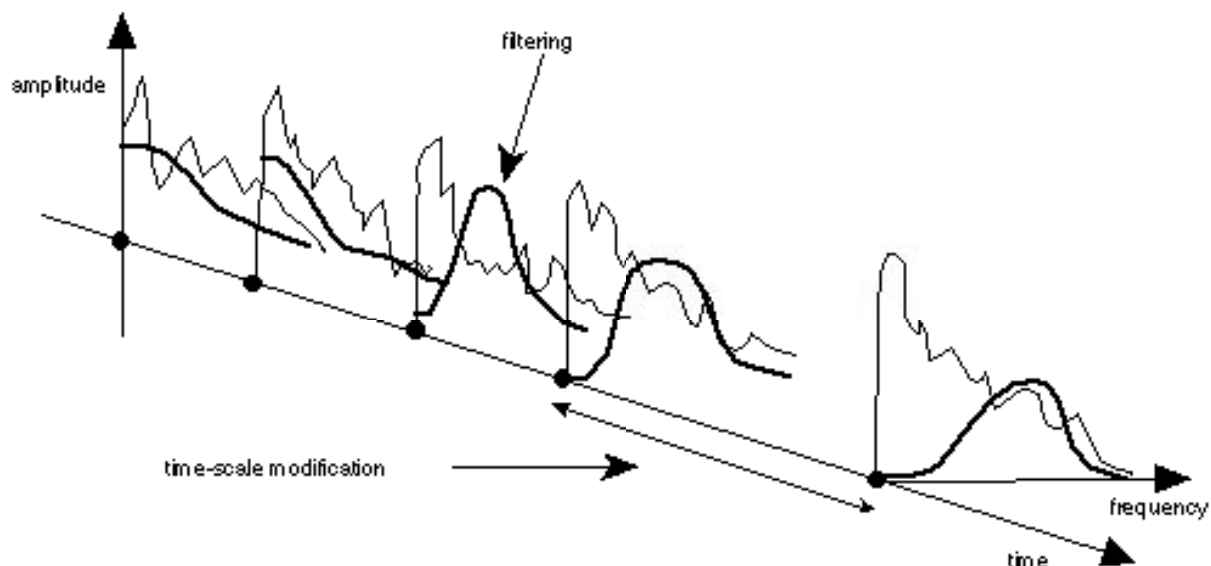
A temporal modification consists of dilating or contracting, at the synthesis stage, the time interval separat-

ing the spectra provided by the analysis; the rate at which the spectrum evolves is thus changed. A sound can thus be expanded or compressed without altering its frequency content. A frequential modification alters the composition of a spectrum at a given moment.

Frequential modifications serve to filter the sound by multiplying the spectrum by the frequency response curve of a filter, or to transpose the sound to a higher or lower frequency. The filter is used to modify the distribution of energy in the spectrum, and thus the timbre of the input sound. As in the example, the filtering can vary during the processing.

AudioSculpt also allows the possibility of processing two sounds at once, and of combining their spectral representations. This combination involves the generated spectra, and thus makes cross-synthesis possible. This option, which can be used in order to generate fairly complex combinations of two sounds, is particularly useful from a musical standpoint.

This illustration shows thge two types of sound modification, temporal and frequential:



Once the series of spectra (possibly modified) is obtained, the resulting sound signal can be synthesized, using the inverse Fourier transform. From the intermediate spectral representation the sound is synthesized and written to a file. As was the case in the forward Fourier algorithm, the inverse Fourier algorithm uses a window to rebuild a segment of the signal.

Each of the operations of filtering, cross-synthesis, expansion/compression, and transposition can be parametrized and controlled as a function of time. This means that blocks of signal can be read and processed separately over time; for example, the frequency response of the filter, the expansion/compression factor and

the coefficients that describe the mixing between two sounds are parameters that the user can specify and vary temporally.

The part on processing is completed with the **Sampling Rate Conversion** function.

General principles

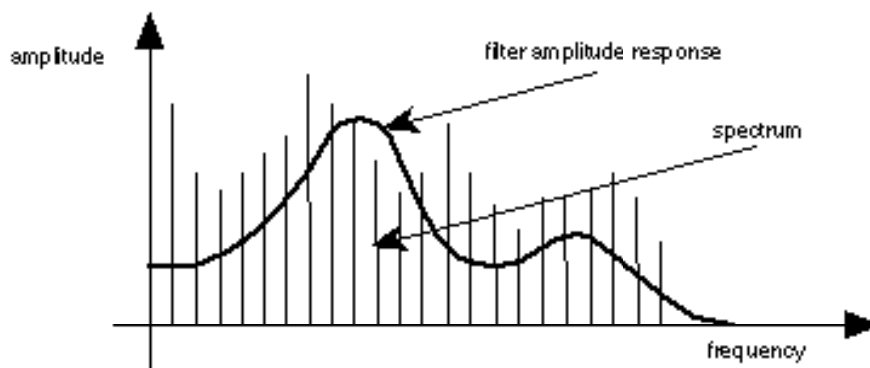
The filtering deals with the spectral information of the sound. The spectral representation that comes from the analysis module is multiplied by the frequency response of a filter. There are several types of filters; that is, there are several ways to define the frequency response curve for a filter. In AudioSculpt you can specify the filter in a graphical way by drawing on the sonogram. Besides the sonogram filter AudioSculpt has four kinds of filtering which operate based on data from a text file: band filtering (alternating stop and pass band filters), breakpoint filtering (filter response curve made up of straight-line segments), surface filtering (filter that is specified as a time-frequency surface) and formant filtering (where the filter envelope is constructed from a combination of second-order filters).

The spectrum of a sound is a series of complex numbers, whose magnitudes correspond to the amplitude spectrum, and whose phases correspond to a phase spectrum. Filtering is the operation of multiplying the amplitude spectrum by an amplitude-frequency curve and of adding a phase-frequency curve to the phase spectrum.

With breakpoint filtering, it is possible to change both the amplitude and the phase spectrum of the input signal. By varying (within bounds) the instantaneous frequency of the components measured in each analysis channel, it is possible to produce spectral distortion in the signal, like, for example, limited transposition of a sound. But this does not allow arbitrary transposition to be performed, because this operation works only with instantaneous frequencies.

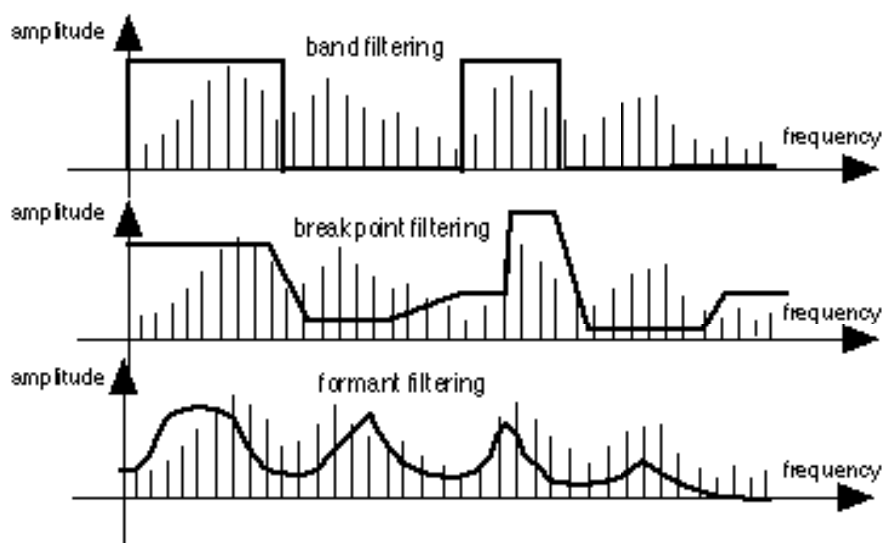
The multiplication of an amplitude spectrum by a curve serves to change the amplitudes that were provided for each analysis channel. The amplitude for each analysis channel is multiplied by a coefficient; if this coefficient is different from 1, the amount of energy in this channel is modified. As with an equalizer, it is possible to attenuate or reinforce certain frequency bands, reinforce or attenuate certain harmonics or entire zones in the spectrum, giving a formant-like effect, etc. In AudioSculpt, several filtering modes are possible,

each characterized by a specific type of filter amplitude response.



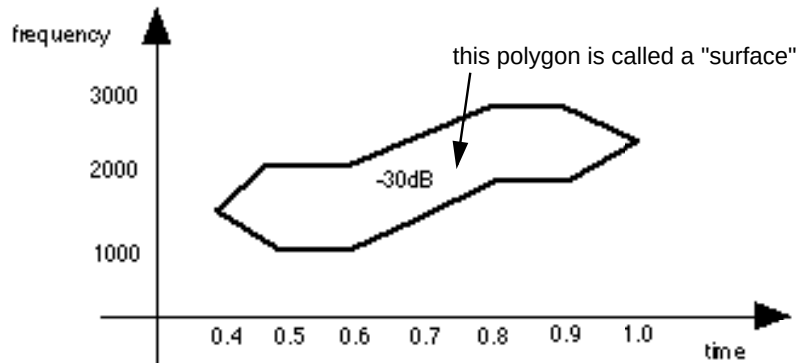
- "band" filtering preserves or removes the energy in a specified band of frequencies.
- "breakpoint" filtering is used to create a response curve that is more general than the band filtering type; the curve is made up of linear segments of variable gain.
- "formant" filtering is useful for creating an envelope with a formant-like structure. One typical use of formant filtering is to apply a succession of formant envelopes (as in articulate speech) to an instrumental sound. This type of filtering is also useful for creating spectral resonances that can be controlled by their center frequencies and bandwidths.

This figure shows these three types of filters in AudioSculpt:



- "surface" filtering is used to modify the gain of a surface defined in the frequency-time plane.

This figure shows an example of surface filtering in AudioSculpt:



The type of filtering can not change during processing, but the parameters that control these various filters (cut-off frequency, gain, bandwidth, etc.) can.

All four types of filters can be specified in AudioSculpt using a text file. AudioSculpt also offers a "sonogram" filter which is a surface filter that can be specified graphically.

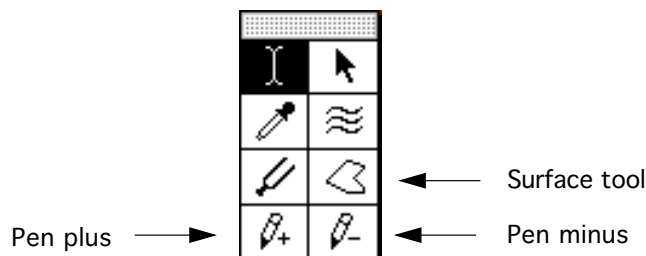
One essential feature of AudioSculpt is that it permits temporal control of filtering. Once the type of filter has been designated, the user can specify any number of filter coefficients that are to take effect at a given moment in the sound; AudioSculpt passes from one such configuration to the next by interpolation (linear interpolation in time). The sound's timbre is modified as a function of the evolution of the filters. This dynamic control of filters can be used not only to bring about timbral evolution, but also to create rhythmic effects by alternating different filter configurations. For example, the rapid alternation between low-pass and high-pass filters on a noisy input sound imposes on the sound a rhythm determined by the speed of the alternation.

The sonogram filter

The sonogram filter is a surface filter, but with AudioSculpt these surfaces can be drawn directly on the sonogram instead of working with text files. Once you have the sonogram displayed you can take the surface tool or a pen in the toolbox and draw the regions that will be filtered on the sonogram. The surface tool is used in a different way than the pens. With the surface tool you draw a surface and then give it a gain. The gain defines how much a region in the sonogram will be amplified or attenuated. With the pens, on the other hand, you can color a region. This region will either be completely cut out of the sonogram (in case of the pen minus) or (in case of the pen plus) this region will be left untouched, but the rest of the sonogram will be eliminated. These kinds of tools are different, which explains why their surfaces are not shown simultaneously on the sonogram. Shifting from a surface tool to a pen will hide the surface tool's regions (they are kept in memory though) and display the previously drawn pen traces. Shifting back to the surface

tool will display its surfaces again and hide the pen's surfaces. Here's a description of how to manipulate these tools.

The surface tool



With the **Surface tool**, you can draw surfaces on the sonogram pane (see also the Quick tour at the beginning of this manual). These surfaces specify the zones of the sonogram that will be filtered.

The surface tool offers two ways of drawing surfaces:

- You can freely draw a line with the mouse holding the mouse key down. When you release the mouse key the surface will be automatically enclosed.
- You can specify the surface as a polygon. Every click with the mouse defines one angle point and one side of the polygon. To close the polygon you must double-click.

You can select and move a surface with the pointer tool. A selected surface has a flashing borderline. To delete a surface you select it and then press the option and delete key simultaneously. With the pointer tool you can also adjust the form of a polygon or rectangular surface by moving its angle points. The label attached to the surface displays the gain that will be applied to that region in the sonogram filter (measured in dB). To change the gain, you select the region and type the number on the keyboard. A positive number means the region will be amplified; a negative number means the zone will be attenuated. In case several surfaces overlap, their gains will not be added; only the gain of the surface on top will be applied. Note that this differs from the Surface Filter (see below).

The pen plus and pen minus

With the pen plus (also called "pass pencil") you can draw freely on the sonogram to mark a region. The regions that are marked with the pen plus are submitted to a band-pass filter, which means that only these regions are left untouched. That part of the sonogram that is left unmarked is eliminated. You could, for example, retain a couple of partials of a complex sound.

The regions that are marked with the pen minus will be eliminated, the rest of the sonogram is left un-

changed. You could for example cut out disturbing elements such as noisy clicks.

Band filtering

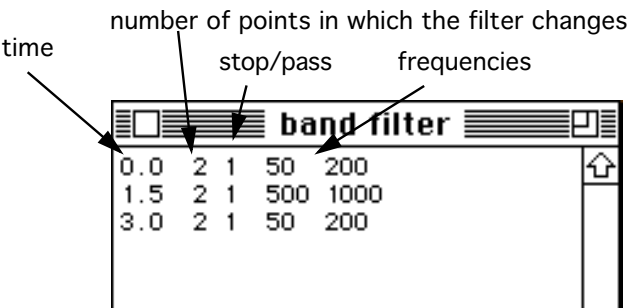
In this mode, a filter is defined as a set of alternating stop and pass bands. The frequencies delimiting these regions can change with time; the number of frequencies used to define the bands and the state of each band can not vary during processing.

The filter parameters are defined from within a text file. You should create (or open) a text file, and enter the Band filter data (also called a parameter table).

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: the number of border frequencies
- Third column: flag specifying if the first band is a stop or pass band
- Following columns: the values of the border frequencies

Take for example the following parameter table:



time	number of points in which the filter changes	stop/pass	frequencies
0.0	2	1	50 200
1.5	2	1	500 1000
3.0	2	1	50 200

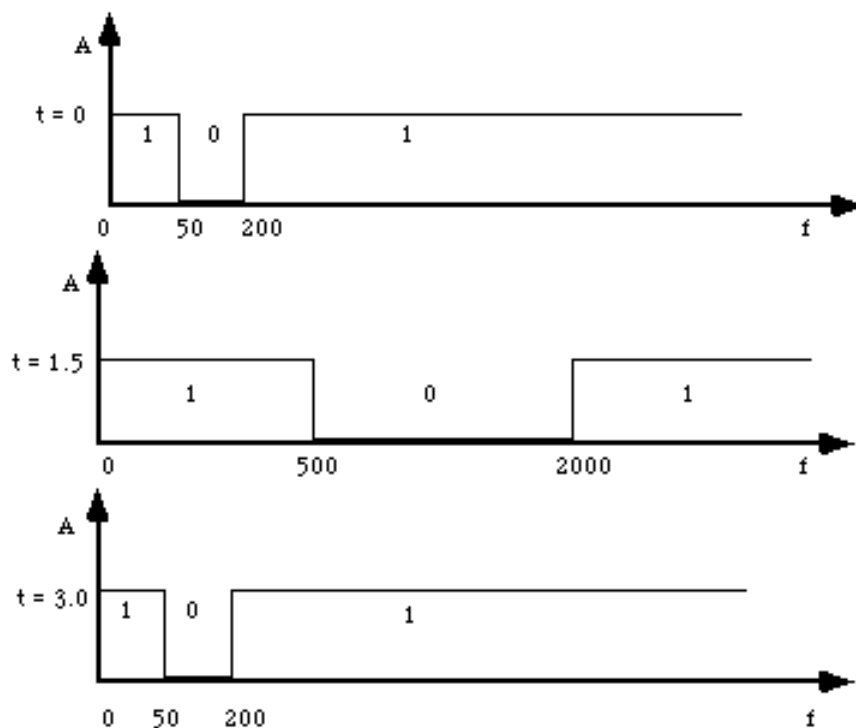
The lines are separated by carriage returns (Return key) and the numbers on each line are separated by blank spaces or tab strokes.

The first column gives the time describing the filter’s time evolution. The second column gives the number of points in which the filter changes, and the third column gives the state of the first band, where :

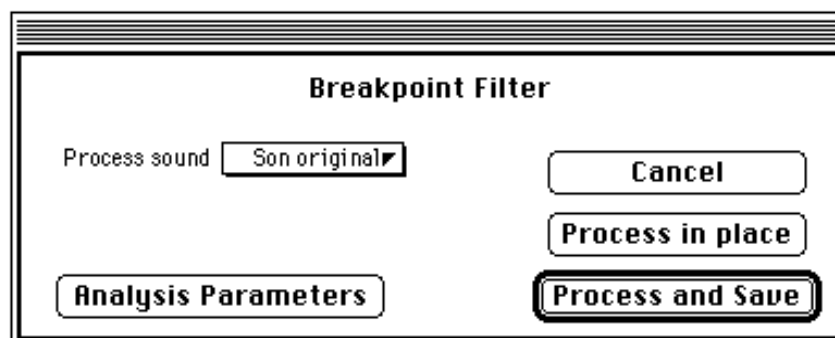
- 0 Stop band;
- 1 Pass band.

Successive bands will alternate between 0 and 1. In this example, there is one stop band. The stop band lies between 50 Hz and 200 Hz at 0 sec, and because AudioSculpt always interpolate linearly the data from one line to the next one, this band moves (linearly) to the frequency region between 500 Hz and 1000 Hz, which

is attained at 1.5 sec, and then returns to its initial state at 3 sec.



Breakpoint filtering



In this type of filtering, the frequency response curve for a filter is defined as a set of points connected by lines. There are three possible modes of operation for this type of filter: amplitude response, phase response,

or amplitude and phase response.

Mode	Operation
0	amplitude response
1	phase response
2	amplitude and phase response

Depending on the mode, the points will be ordered pairs of the form (frequency, amplitude), (frequency, phase), or triplets of the form (frequency, amplitude, phase). Amplitudes are given in decibels, phases in radians, and frequencies in Hertz.

The following parameter file specifies breakpoint filtering that passes from a configuration in which the extreme regions of the spectrum (fundamental and high harmonics) are attenuated to the inverse configuration at 1.5 sec and then comes back to the initial filter at 3 sec. Only the amplitude response is specified in this filter (mode 0), and the filter is defined by amplitude-frequency pairs.

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: the mode of the filter (0, 1 or 2)
- Third column: the number of breakpoints
- Following columns: the coordinates of each breakpoint (frequency-amplitude for mode 0, frequency-phase for mode 1, frequency-amplitude-phase for mode 2).

mode
here: 0, amplitude response

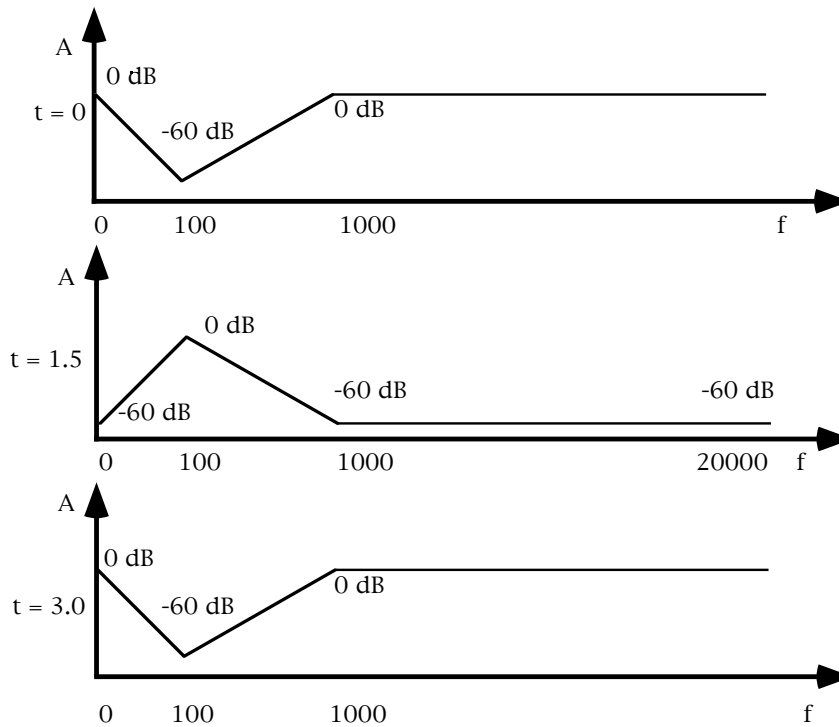
number
of pairs

frequency/amplitude pairs

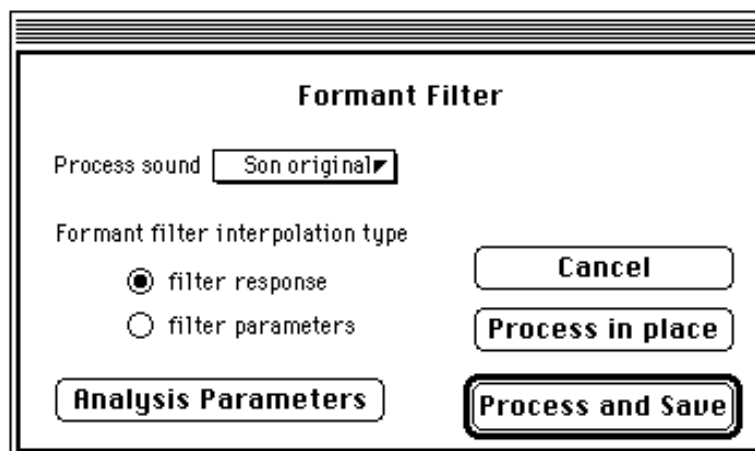
time

breakpoint filter										
0.0	0	3	0	0	100	-60	1000	0		
1.5	0	4	0	-60	100	0	1000	-60	20000	-60
3.0	0	3	0	0	100	-60	1000	0		

In this type of filtering operation, the number of pairs defining the frequency response curve for the filter can change in time. They change linearly from one line to the next. The mode of operation, however (amplitude, phase or amplitude and phase) cannot change.

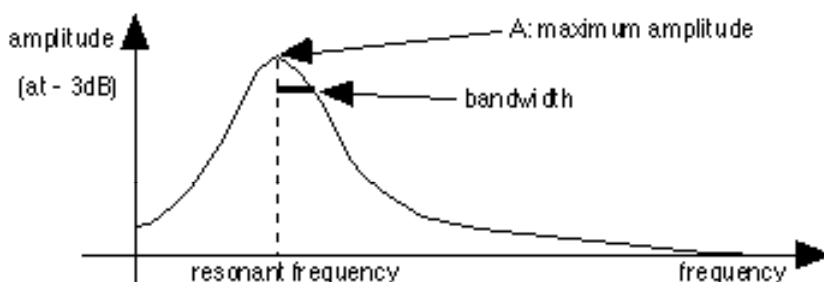


Formant filtering



In this mode, a filter is defined as the sum of a certain number of formants (These are technically called "2-

pole recursive filters"), each of which is described by a central frequency (in Hertz), a gain (in decibels) and a bandwidth (in Hertz). The gain corresponds to the height of the maximum for the curve. The resonant frequency is the frequency at which the curve attains its maximum, and the bandwidth corresponds to the half-width of the curve when the amplitude is 3 dB below the peak value. This figure formant filter response curve (2-pole recursive filter):



There are two types of formant filtering: one is called "Filter response", the other is called "Filter parameters".

Filter response	the number of formants can vary with time. This mode is called "fifof filtering" in SVP.
-----------------	--

Filter parameters	the number of formants is fixed. This mode is called "fof ¹ filtering" in SVP.
-------------------	---

When you launch the command **Formant filtering**, the dialog box allows you to choose between the two types of formant filtering interpolation.

You write into a text file one or several lines in which you specify the time, the number of formants, and, for each formant, its frequency, amplitude, and bandwidth (in Hz), as described hereafter.

- First column: time at which to apply the parameters
- Second column: the number of formants
- Following columns: the central frequency, the gain, and the bandwidth of each formant

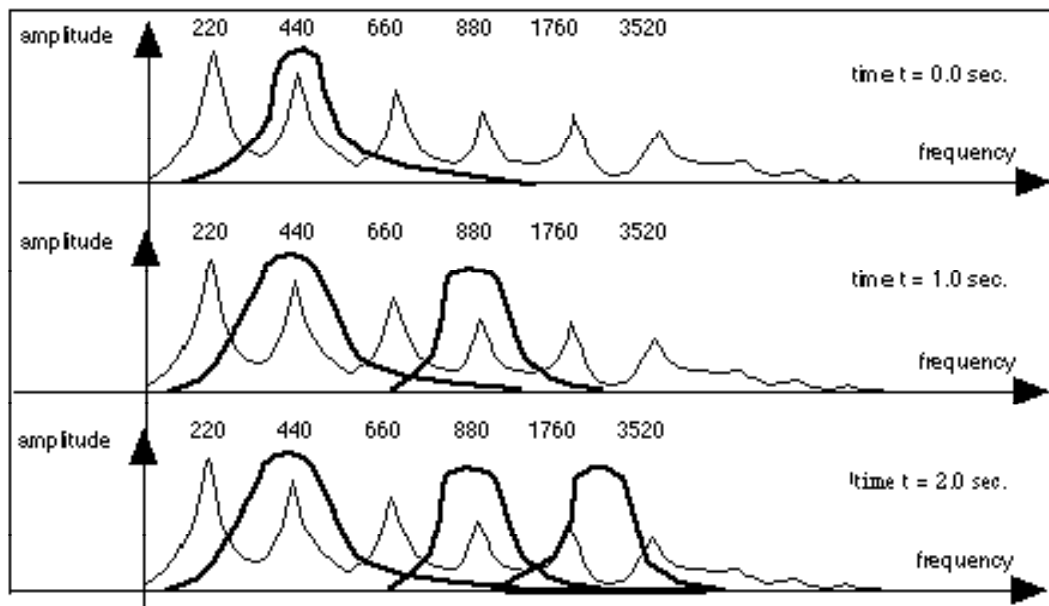
The following file defines a formant filtering that applies a spectral envelope with an increasing number of formants (i.e. it can only be used in "response" mode):

1. "fof" stands for "*Fonction d'onde formantique*". This is an adaptation from a work by X. Rodet.

formant filter parameters									
0.0	1	440	0	50					
1.0	2	440	0	50	880	-2	100		
2.0	3	440	0	50	880	-2	100	1760	-2 200

The first column describes the time, the second column the number of formants and then for each formant the central frequency, the gain and the bandwidth are given.

The following figure shows the evolution of the envelope applied to the sound's spectrum. The formants are represented individually, before they are added.



When you ask the **Formant Filter** in the **Processing** menu, the dialog will ask you for the soundfile to which you want to apply the processing.

Surface filtering

The surface filter modifies the amplitude of a surface defined in the amplitude-frequency plane. Every sur-

face filter modifies the sound according to a fixed gain. In the re-synthesis, the spectral components that fall in the regions defined by the surface will be amplified or attenuated according to the gain attributed to the surface. The surfaces can overlap: in that case their gains will be added. The command **Sonogram Filter** (see above) offers a very similar filter that can be operated graphically.

The format of the parameter file is as follows:

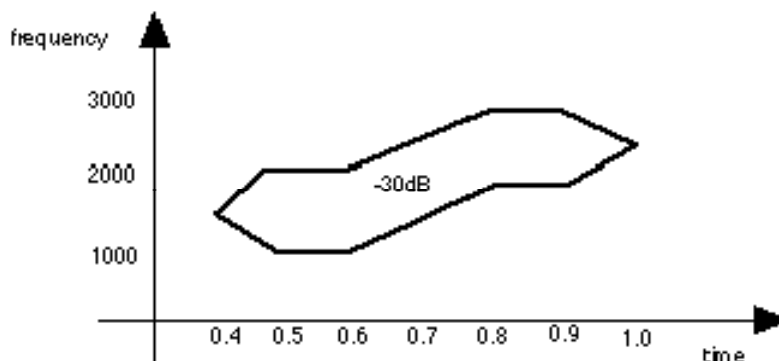
- First line: start time, number of points, end time and gain
- Following lines: time, lower and upper border of the surface

The following example describes a surface filtering:

surface filter parameters			
0.0	4	5.0	-40
1.0	1000.	2000.	
2.0	1000.	2000.	
3.0	2000.	3000.	
4.0	2000.	3000.	

The first line gives the starting time of the surface, the number of frequency couples describing the surface, the end time of the surface and the gain. The following lines read the time of the frequency points and then two frequencies describing the lower and upper border of the surface.

The following figure shows another example of a surface, displayed graphically this time. In this figure, an interpolation is done between the start time and the time of the two first points as well as between the time of the last two points and the end time.



With AudioSculpt new sounds can be synthesized based on the analysis of one or two existing sounds. Cross synthesis is the combination of two sounds at the level of the spectral representation. It is used in order to create different types of mixes between two sounds, hybrid effects, overlap, etc. AudioSculpt permits two types of cross-synthesis: cross synthesis source-filter and generalized cross synthesis

The two sounds' parameters (window size and window step) can also differ according to the input channel. If the window steps are different for each input, the time scales of the spectral representations will be different as well. Since during the synthesis stage, the time scale of the primary input is used as a reference, the time scale of the secondary input will be automatically changed, giving a time expansion/compression. For example, if sound 1 has a window step of 256 samples, then the time interval between the short-term spectra is equally 256 samples. If sound 2 has a window step of 128 samples then the interval between the spectra is 128 samples. In the cross synthesis, the first short-term spectrum of sound 1 is combined with the short-term first spectrum of sound 2, the second spectrum is combined with the second spectrum, and so on... Since the spectra of sound 2 has a time interval that was initially half the time interval of sound 1, the spectral data of sound 2 will be stretched by a factor 2. This corresponds to a dilation by a factor 2.

The cross synthesis source filter

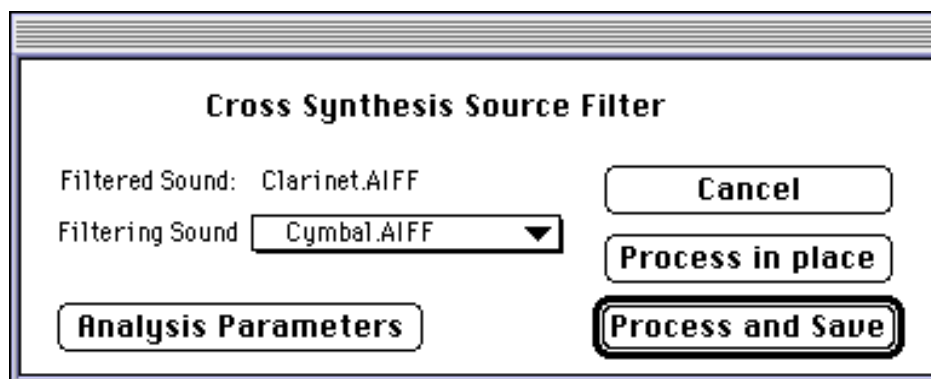
Cross synthesis source filter consists of multiplying the FFT spectrum from the primary input by a spectral envelope resulting from linear predictive analysis on the secondary input. The first sound is thus filtered by the spectral envelope of the second. This type of operation is very close to the source-filter synthesis model; in AudioSculpt, the source is the first sound and the filter is the spectral envelope of the second. A common application is the filtering of an instrumental sound by the voice, so as to give it a “speaking instrument” effect.

One of the sounds is analyzed by the FFT and the other by linear prediction (LPC). Two parallel spectral representations are obtained, one being a series of FFT spectra, the other a series of LPC spectral envelopes.

The sequences are combined by multiplication of the FFT spectra by the LPC spectral envelopes.

This cross-synthesis is called source filter because it is similar to a source filter synthesis system in which the output signal is calculated by passing the source signal through a time-varying linear filter. Such a system models the voice (glottis and vocal tract) and certain instrumental sounds. In this voice model, the source signal corresponds to an excitation, and the linear filter to a resonator. It can either be a periodic impulse train or a noise source. The resonator is characterized by its resonant frequencies, also called formants. The resulting sound spectrum is the product of the spectrum of the source signal (harmonic or noise spectrum) with the frequency response curve of the resonator (curve corresponding to the set of formants). In AudioSculpt the signal source is a sound, and the resonator is another sound. For this operation, only the spectral envelope for the filtering sound is evaluated.

The applications of source-filter cross-synthesis are numerous, and are, in general, quite interesting. One example, now commonplace, is the filtering of an instrumental sound by speech, or the inverse. In the first case, one hears the instrumental sound filtered by the vocal tract, shaped, as it were, by the phonemes of speech. The instrument's timbre is deformed by the vocal formants. In the second case, the vocal impulse passes through the resonating instrument body. The vocal timbre is determined by the instrument's coloration.



The **Cross Synthesis Source Filter** command in the processing menu brings up the dialog displayed below. Filtered sound is the sound which will be filtered with the spectral envelopes of the filtering sound. In this example, the spectral envelope of the Cymbal.AIFF is applied to the Clarinet.AIFF. To set the parameters of the analysis, click the **Analysis Parameters** button. The general parameters for both channels can be set (window size, window step, analysis window) as well as the number of poles in the LPC analysis, the size of the FFT and the resynthesis window. Remember that the higher the number of poles, the more detailed the spectral envelope. You may want to use a few poles when you only want to model the most important formants of the spectrum envelope.

Analysis Parameters

	Channel 1	Channel 2
Window size	1024	1024
Fund. Frequency	215.3	215.3

Window step	128.000	128.000
Analysis Window	Hamming	Hamming
		1pc order 30
FFT Size	1024	
Synthesis Window	Hamming	

Factory Settings

Cancel

OK

Generalized cross synthesis

Generalized cross synthesis is an extension of the previous cross synthesis. The spectra of the two sounds (initially represented by their real and imaginary parts) are decomposed into amplitude and phase spectra, which can be manipulated independently of one another. This independent control over amplitude and phase gives a considerable variety of sound combinations. One particularly attractive application is making hybrid sounds by applying the phase spectrum of one sound to the amplitude spectrum of another sound or “crossfading”, that is the progressive transformation of the amplitudes and phases of one sound into the amplitudes and phases of another.

Here is a technical description of the operation.

For each analysis window, the complex spectrum is broken up into an amplitude spectrum and a phase spectrum. The phase values are transformed into instantaneous frequencies. If A_1 and ϕ_1 are the amplitude and phase spectra for a particular window, and A_2 and ϕ_2 are amplitude and phase spectra for a second sound, at the same location, then the amplitude and instantaneous phase for the synthesized frame are obtained using the following formula:

$$A = XA_1 + xA_2 + qA_1A_2$$

$$\phi = Y\phi_1 + y\phi_2$$

The formula basically states that the amplitude spectrum for the output signal is a linear combination of the amplitude spectra of the input signals, augmented by a co-multiplicative factor scaled by the coefficient q. Given independent control over amplitude and instantaneous frequency, it is possible to alter the frequencies of a sound's partials without changing their amplitudes, and vice versa.

We can obtain a hybrid sound with the combination (X = 1, x = 0, Y = 0, y = 1, q = 0). Here the instantaneous frequency spectrum of the secondary input is applied to the primary input sound. With the combination (X = 0, x = 1, Y = 1, y = 0, q = 0) the amplitude spectrum of the secondary input is applied to the primary input sound.

The multiplication of the two amplitude spectra is a rather delicate matter. When this operation is used alone (X, x are zero), the resultant spectrum may be zero if the two spectra are complementary; that is, when their spectra do not overlap at all.

If the coefficients X, x, Y, y and q are constant, they can be specified in the dialog when choosing the **General Cross Synthesis** command; if not, they can be provided in table form from a text file.

The following example creates a non time-varying combination: the phase spectrum of Cymbal.AIFF secondary input is applied to the amplitude spectrum of Clarinet.AIFF.

Generalized Constant Cross Synthesis

Sound 1

Clarinet.AIFF

Sound 2

Cymbal.AIFF

amplitude cofactor

0.0

amplitude scale factors

1.0

0.0

phase scale factors

0.0

1.0

Analysis Parameters

Cancel

Process in place

Process and Save

In the example below, the parameters are specified in a text file:

general cross synthesis					
0.0	1.0	0.0	1.0	0.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0
2.0	1.0	0.0	1.0	0.0	0.0
3.0	0.5	0.5	0.5	0.5	0.5
	X	x	Y	y	q

The first column specifies the time of the parameters and on each line X, x, Y, y, q for that time are given. In the above example, the amplitude and phase spectra of the output vary linearly between the two sounds, between 0 and 1 second.

Between 1 and 2 seconds, the output returns to the first sound, and after 2 seconds the resulting sound is a combination of the two amplitude and phase spectra and of the multiplication spectrum.

Depending on the input, generalized cross-synthesis can create transitions that are quite complicated, in terms of timbre. Timbral interpolation between two different and contrasting sounds, like, for example, a pitched sound and a noise sound, is especially interesting.

Time scale modification

Sound expansion and compression are techniques used in order to change the rate at which events occur, in order to change the duration of a sound while retaining its character, or to change both speed and duration simultaneously. This expansion/compression allows the modification of the time scale of events without changing the spectral information they contain.

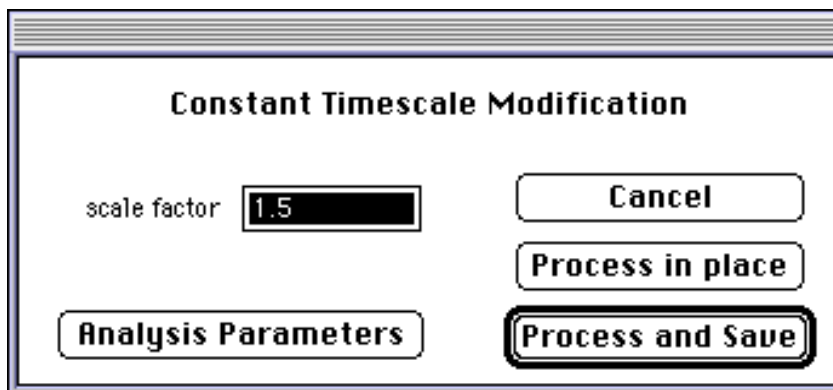
Changing the speed at which a sound evolves makes it easier to perceive the details of the sound. For example, the expansion of a piano sound allows the striking of the hammer on the strings and the resonance of the vibrating string to be isolated. In the same way, slowing down a recording of speech makes it easier to examine the articulation of the phonemes. This type of processing provides a useful effect by deforming the time-scale of the micro-events which make up a sound.

Changing the duration of a sound is often necessary when attempting to synchronize digital sounds, or when one needs to insert a sound of a pre-determined duration into a piece of music. Since it is impractical, and sometimes impossible to sample a sound of exactly the desired duration (whether the sampling is done from a recording, or from a live player), one must be able to change this duration arbitrarily, without degrading the quality of the sound too much.

AudioSculpt offers a constant time scale modification: the argument is then a coefficient referring to a constant stretch factor. You can also obtain a time-varying time scale modification when you use a file of coefficients. In both cases, a coefficient bigger than 1 increases the length of the sound, whereas a coefficient between 0 and 1 reduces it: the duration of the sound is multiplied with the scaling factor. To obtain a good sound quality it is important to set the analysis parameters properly.

Constant time scale modification

You can modify the length of a soundfile with a constant factor when choosing the Time Scale Modification command in the Processing menu. In the dialog box below, next to **scale factor** you can type the factor to be used. In this example, the time scale is multiplied by 1.5, and the resulting soundfile is thus 50% longer than the original.



Similarly, to reduce the total duration of the sound, type for example 0.5. The resulting sound will then be half as long as the original one.

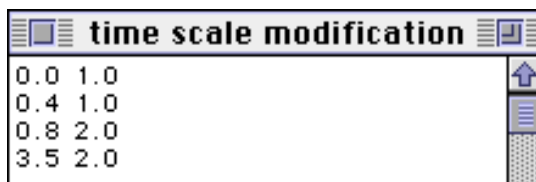
Time-varying time scale modification

It can be useful to have varying expansion or compression. For example, the attack of a percussive sound is left un-modified, while the rest of the sound, the stable part, is stretched. The scaling parameters must be typed in a text file. The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: scale factor

On each line, these values should be separated by at least one space or Tab. Each line should be terminated by a carriage return (use the Return key).

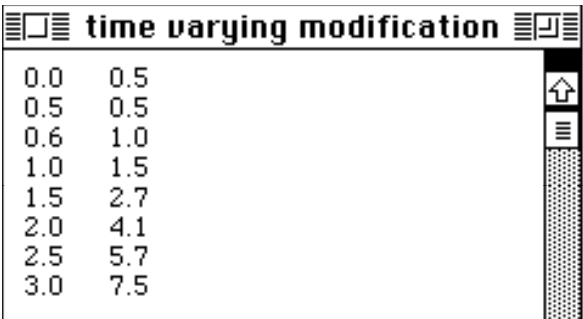
To expand a sound without modifying the attack portion, type:



In this example, time stretching will begin at 0.4 seconds, just after the end of the attack. The scale factor varies from 1 to 2 between 0.4 and 0.8 seconds and stays equal to 2 after that. Between 0 and 0.5 seconds,

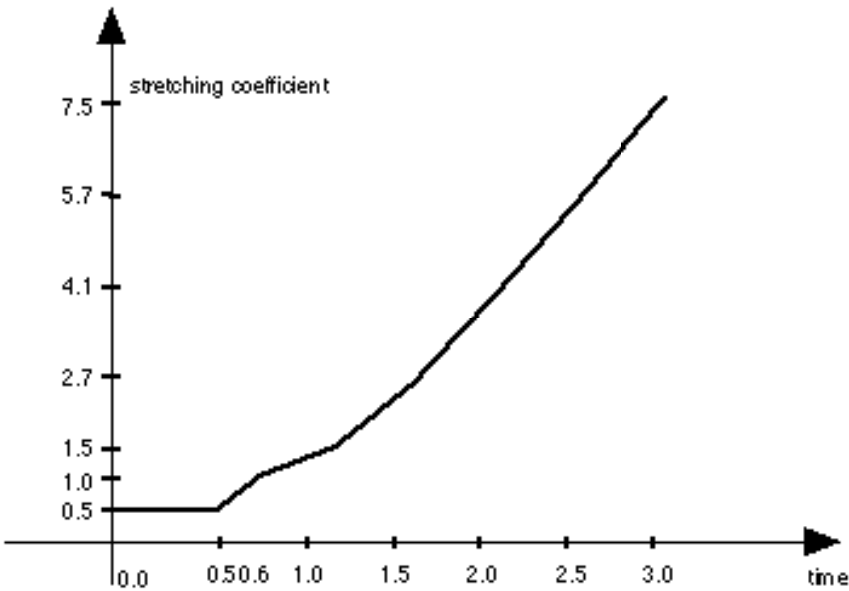
the coefficient is equal to 1, and hence there is no stretching over this segment. When the scaling parameter changes too abruptly, the resulting soundfile might be corrupted.

To compress the attack, and still expand the rest of the sound, this time by a factor that increases with time, type:



0.0	0.5
0.5	0.5
0.6	1.0
1.0	1.5
1.5	2.7
2.0	4.1
2.5	5.7
3.0	7.5

AudioSculpt performs a linear interpolation in time between the coefficients given on each line. For example, between 1.0 and 1.5 seconds, the stretching factor changes linearly from 1.5 to 2.7 (an increase of 1.2). Between 1.5 and 2.0 seconds, the coefficient changes linearly from 2.7 to 4.1 (an increase of 1.4). From 2.0 to 2.5 the coefficient increases by 1.6 etc. It is apparent that the stretching factor increases more and more rapidly with time, and thus the sound is slowed down by more and more. Here is a figure of the stretching coefficient as a function of time



Improving the quality of a time-stretched sound

If the spectral analysis is lacking in resolution, artifacts can appear, simply because the time stretching operation done by AudioSculpt uses the spectral representation of the sound. If the short-term Fourier transform is not very selective, the phase measurements will not be reliable, and distortions will occur when the sound is resynthesized.

If the quality of a time-stretched sound is not sufficient (a bad time-stretched sound sometimes has a metallic or flanged character), the problem may be the lack of resolution in the spectral analysis. In this case, the window size, which is 1024 samples by default (or 23.2 milliseconds at 44100 Hertz), should be increased as much as possible. Suppose that the input sound is periodic, with a fundamental of 220 Hz and a sampling rate of 44100 Hz. The window size should be equal to 4 or 5 times the period of the sound, or $4 \times (44100 / 220)$ samples (approximately 800 samples). In AudioSculpt, five times the period length is selected as the window length (see Part 2 Analysis: FFT)

Transposition

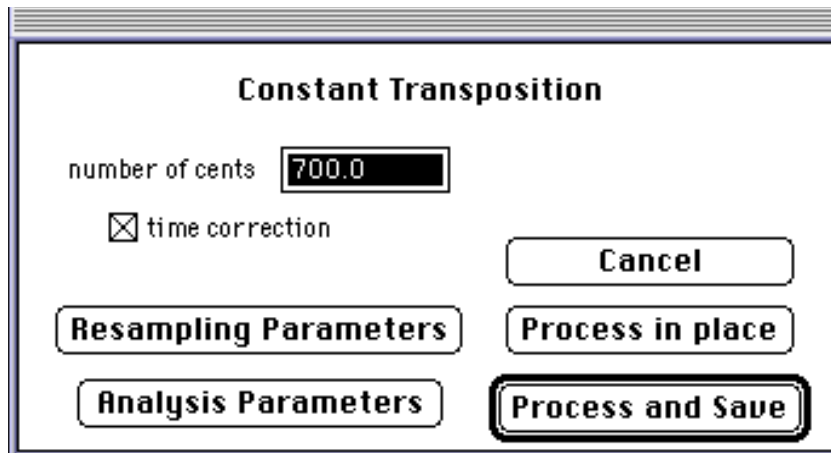
When creating new sounds one might want to shift a recorded sound to a higher or lower frequency. The whole sound can be transposed with a constant interval, or the interval can vary in time. This way it also possible to adjust slight frequency deviations of the sound according to the desired frequency. AudioSculpt offers a constant transposition and a time varying transposition. Both types of processing specify the transposition interval in cents. The cent unit stands closer to musical notation than Hertz. The next paragraph tells more about it.

Cents scale

When we divide the interval of an octave into 1200 equal steps (**footnote:** equal here means according to the frequency ratios: the distance between f_3 and f_2 is equal to the distance between f_2 and f_1 if f_3/f_2 equal to f_2/f_1), one step corresponds to one cent. If two frequencies are one cent apart their frequencies have a ratio of $2^{1/1200}$. A distance of 100 cents corresponds to a half-step, 200 cents corresponds to 1 whole step and so forth.

Constant transposition

For a transposition with a constant interval choose the Transposition function in the processing menu. A dialog box is displayed. In it, you can specify the transposition in cents. In this example the sound is transposed by 700 cents, or a perfect fifth.



If the time correction option is marked, AudioSculpt will adjust the length of the new soundfile to make it just as long as the original soundfile. The time adjustment requires a time scale modification. AudioSculpt will automatically calculate the time scale factor. In case of a time correction, it is necessary to set the analysis parameters.

Time-varying transposition

A time-varying pitch transposition is possible if the active window is a parameter file describing the evolution of the transposition.

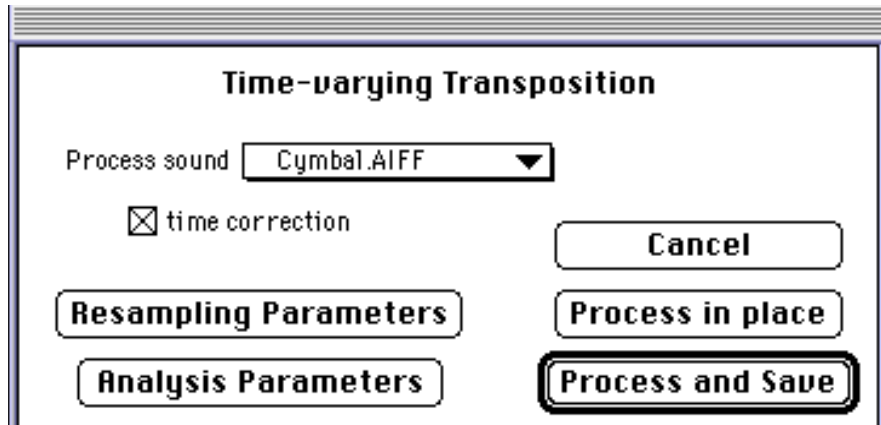
The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: interval of the transposition in cents

Here is an example of such a parameter file:

time-varying transposition	
0.0	0
1.0	100
2.0	0

To run the transposition choose the **Transposition** command. A dialog box will appear. In the menu besides **Process sound**, you can specify the sound on which to apply the transposition.



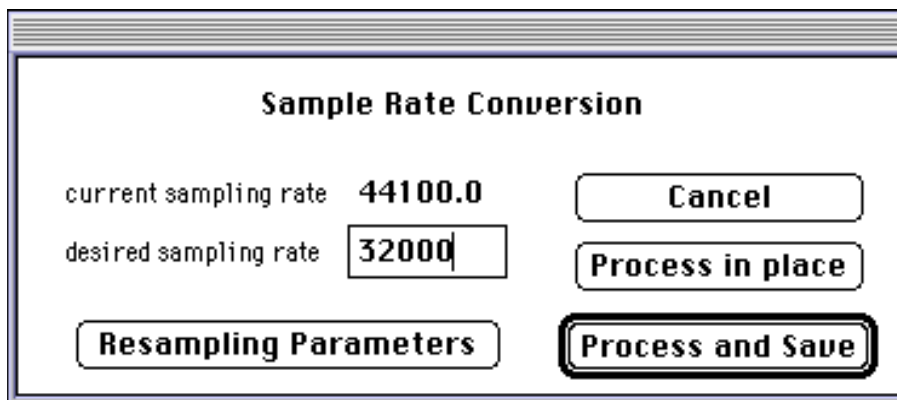
If you check the time correction checkbox (as in the illustration above), AudioSculpt will adjust the duration of the output sound so that it matches the original's sound duration. You'll see that when you select time correction, the **Analysis Parameters** dialog box becomes available. It's a good idea to open it and set the parameters according to your sound's characteristics.

Sample rate conversion

Most often sound is sampled at 44100 Hz. This is the sampling rate used for Compact Discs for example. Other sampling rate standards are 32000 Hz and 48000 Hz.

The sampling rate is mainly determined by the highest frequency in the sound. The half of the sampling rate is called the Nyquist frequency. A sampled sound should not contain any frequency components with a frequency higher than the Nyquist frequency. So if your sound is sampled at 44100 Hz, the highest frequency component will be 22500 Hz. In other words, if the sound has no components above, let's say 10000 Hz, then you can sample the sound at 20000 Hz, or, given a ten percent margin, 22000 Hz, without degrading the sound quality. Moreover, sampling at a lower rate can save you space on the hard disk.

You can change the sampling rate of the soundfile with the **Sample Rate Conversion** command in the **Processing** menu. You type the new sampling rate in the text zone beside **desired sampling rate**. Using the **Resampling Parameters** button, you can set the length of the interpolation function measured in samples.



Part 4

Menu reference

1 The File menu

The **File** menu is composed of the standard file management functions:

New

The **New** command creates a new text document.

Open...

The **Open** command lets you open a soundfile, a sonogram or a text file. The sound formats that AudioSculpt recognizes are the AudioSculpt, AIFF and Sound Designer format. The size of text files are limited to 20,000 characters. If you open a sonogram analysis, the corresponding soundfile will be opened as well.

Close

The **Close** command closes the active window (text, sound or sonogram window). If the document has unsaved changes, AudioSculpt asks you if you want the changes to be saved.

Save, Save As...

The **Save** and **Save As** commands lets you save a soundfile, a sonogram or a text file.

Set Default Folders

The **Set Default Folders** option is important only when you use the command line. It lets you specify where AudioSculpt has to look for the sound and parameter files, and where the console files generated by AudioSculpt will be saved. By default, AudioSculpt takes the folder where the application resides. After you have specified the folders, they will be taken as the default folders and will be memorized by AudioSculpt. The three folders can be the same folder, and can be changed at any time.

Page Setup...

The **Page Setup** command brings up the standard printer Page Setup dialog box. It lets you set the format of the document to be printed.

Print...

The **Print** command starts printing the document. AudioSculpt permits the printing of the sound and the sonogram image and also of the text files. If the active window only displays a sound pane, then the sound curve will be printed; if the active window displays the sonogram, then the sonogram without the sound curve will be printed. To have a good reproduction it is important to mark the Colors / Grey Scale option.

Quit

The **Quit** command ends the AudioSculpt session. Before closing the application a dialog will ask you if you want to save the last changes made to the text, sound or sonogram.

Undo

The Undo command un-does the last command or key stroke executed. This function is not available for sounds or sonograms.

Cut

The **Cut** command clears the selected text and saves it on the clipboard. This function is not available for sounds or sonograms.

Copy

The **Copy** command saves the selected text on the clipboard without clearing it. This function is not available for sounds or sonograms.

Paste

The **Paste** Command inserts the text copied with one of the previous functions. This function is not available for sounds or sonograms.

Clear

The **Clear** command clears the selected text without saving it. This function is not available for sounds or sonograms.

Select All

The **Select All** command selects all of the text or sound in the active window.

Preferences

Under **Preferences** you can set certain options:

- Display splash-screen at start-up: if the option is checked a window describing the application will appear on the screen when you launch AudioSculpt.
- Automatically save consoles when running scripts: when the option is checked the console windows will be automatically saved in the default console folder (see the Set Default Folders command).
- Hide console window unless using command: if this option is checked, AudioSculpt will not display the console window when a process is activated in the Processing or Analysis menu. The console window is always shown when the command is launched from the command line.

Show Clipboard

The **Show Clipboard** displays the text that was copied last with the cut or copy function.

The **Processing** menu groups together the signal treatment functions.

Time Scale Modification

With the **Time Scale Modification** command you can expand or compress the soundfile in the active window. The expansion factor can be introduced in the text zone next to ‘scale factor’. The duration of the sound is multiplied by the scale factor. If the factor is bigger than 1 the sound will be expanded; if the factor is smaller than 1 the sound will be compressed. The scale factor is only limited by the amount of free space on your hard disk. The re-scaling process uses an FFT algorithm so it’s important to specify the analysis parameters.

A time varying time scale modification is possible if you have a parameter file open. The parameters describe the evolution of the expansion or compression in time. You can choose the soundfile on which to apply the modification in the dialog box.

Sample Rate Conversion

With the **Sample Rate Conversion** command you can modify the current sampling rate of the sound. In the text zone next to **desired sampling rate** you can fill in the new sampling rate. The conversion does not alter the length in time nor the pitch of the original sound. Only the number of samples changes.

The size of the interpolation function can be set using the **Resampling Parameters** button.

Transposition

The Transposition function transposes the sound to a higher or lower pitch. In the text zone you can specify the transposition in cents. If the time correction option is marked, AudioSculpt will adjust the length of the new soundfile to make just as long as the original soundfile. The time scale factor will be calculated automatically by AudioSculpt. In case of a time correction, it is necessary to set the analysis parameters.

A time varying pitch transposition is also possible if the active window is a parameter file describing the evolution of the transposition. In the menu besides **Process sound** you can specify the sound on which to apply the transposition. The option **time correction** is also valid in case of a time varying transposition.

Breakpoint Filter

This functions is only accessible if a parameter file is active.

The **Breakpoint Filter** is a filter which has a transfer function that is a correction of connected linear segments. Three types of transfer functions can be defined: amplitude, phase or amplitude and phase transfer function. Each type of transfer function defines a mode:

Mode	Transfer function
0	amplitude
1	phase
2	amplitude and phase

According to the mode, the points of the breakpoint function are defined as frequency/amplitude pairs, as frequency/phase pairs or as frequency/amplitude-phase triplets. The amplitudes are measured in dB, the phases in radians and the frequencies in Hertz.

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: the mode of the filter (0, 1 or 2)
- Third column: the number of breakpoints
- Following columns: the coordinates of each breakpoint (frequency-amplitude for mode 0, frequency-phase for mode 1, frequency-amplitude-phase for mode 2)

The dialog box allows you to choose the sound on which to apply the filtering.

Band Filter

This functions is only accessible if a parameter file is active.

The **Band Filter** eliminates the spectral components that fall into a stop band and does not affect the components that fall into a pass band. These bands are defined by the frequencies that border them and whether the first one is a pass band (set the flag to 0) or a stop band (set the flag to 1). The nature of the following bands are specified automatically since pass and the stop bands alternate. The values of the border frequencies can vary in time, but the number of bands in the filter and their mode (pass or stop) cannot change.

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: the number of border frequencies
- Third column: flag specifying if the first band is a stop or pass band
- Following columns: the values of the border frequencies

The dialog box allows you to choose the sound to which to apply the filtering.

Surface Filter

This function is only accessible if a parameter file is active.

The **Surface Filter** modifies the amplitude of the spectral components that fall into a surface defined in the amplitude-frequency plane. Every surface filter has a fixed gain. The surfaces can overlap: in that case their gains will be added. The command Spectrogram Filter offers a very similar filter that can be operated graphically.

The format of the parameter file is as follows:

- First line: start time, number of points, end time and gain
- Following lines: time, lower and upper border of the surface

An interpolation is done between the start time and the time of the two first points as well as between the time of the last two points and the end time. The dialog box allows you to choose the sound on which to apply the filtering.

Formant Filter

This function is only accessible if a parameter file is active.

The **Formant Filter** is a filter composed of second order filters. Each filter is defined by three parameters: the central frequency (in Hertz), the gain (in dB) and the band width (in Hertz). Every second order filter corresponds to a formant. It is possible to choose the kind of interpolation that is used in calculating the temporal evolution of the filter. The first type interpolates the filter responses (the filter responses are calculated by adding the second order filter responses). The second type interpolates the filter parameters (central frequency, amplitude, bandwidth). The type of interpolation can be defined in the dialog box. If you choose the filter response mode, it is possible to vary the number of formants in time. In the filter parameter mode, it is impossible to change the number of formants.

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- Second column: the number of formants
- Following columns: the central frequency, the gain, and the bandwidth of each formant

The dialog box allows you to choose the sound on which to apply the filtering.

Sonogram Filter

With the **Sonogram Filter** you can filter the sound in a graphical manner. The spectral surface of the sound must be calculated with the **Sonogram Analysis** command in the **Analysis** menu (see also Part 2 Analysis and the introduction). The surface tool and the pencil plus and minus allow you to draw a surface on the sonogram. If the surface is drawn with a surface tool, a gain can be given to it by selecting the surface and typing the value of the gain on the keyboard. A positive number corresponds to an amplification; a negative number corresponds to an attenuation. If the surface is drawn with the one of the pencils, the colored area will be cut out (in case of the pen minus) or nothing but the colored area will be kept (in case of the pen plus). The filter will be applied to the actual soundfile and the resulting sound is saved in a new soundfile. This command is not available with the command line.

Cross Synthesis Source Filter

In the **Cross Synthesis Source Filter** the sound marked next to 'Filtered Sound' will be filtered with the spectral envelope of the sound selected under 'Filtering Sound'. The spectral envelope of the Filtering Sound is calculated using the linear prediction analysis. The parameters of the FFT analysis and of the LPC analysis can be set with the **Analysis Parameters** option.

Generalized Cross Synthesis

The **Generalized Cross Synthesis** is a synthesis technique that for example permits a gradual spectral transition between two sounds. The short-term spectra of both sounds are decomposed in an amplitude spectrum and a phase spectrum. Both spectrums can be manipulated separately. A gradual spectral transition is possible if you use a parameter file describing the temporal evolution of the parameters. The initial soundfiles should be both stereo or mono; the result file has the same configuration as the initial files.

The format of the parameter file is as follows:

- First column: time at which to apply the parameters
- 2nd, 3th, 4th, 5th and 6th column: the values of X, x, Y, y and q

Use Command Line

With the **Use Command Line** function you have access to a box that lets you enter a UNIX-style command line. With this command line you can launch analysis or processing and change its parameters without having to manipulate the graphical interface. The command lines can be saved in a text file so you can create a personal data base of command lines. The lines can then be copied from the text file into the dialog box. The **Process** button will start the analysis. The command line window can stay on the screen (inactivated) while you are working with other files. The text in the dialog box can be copied and cut into a text document such as a script (see **Run Script** below).

The commands for the spectrogram filter and the sonogram analysis are not available through the command line.

In the annexes A and B you will find a complete list of the commands and their options.

A command can also be sent from within a text file. Type the command in a AudioSculpt text file (or open a saved text file) and press the Enter key.

Run Script

The **Run Script** lets you launch a list of processing tasks using the principle of the command line. Different kinds of processing are possible and several soundfiles can be manipulated with one script. Type the command line in a text file (an AudioSculpt text file, choose **New Text Document** in the **File** menu). Each line must be separated by an empty line. Choosing the **Run Script** command will start every process one after another. If you choose Stop in the dialog box that appears on the screen, the current process will be terminated and AudioSculpt will jump to the next command. The **Stop Script** button will stop all commands. You can also run the script by pressing the enter button when the text window is active. You should give different names to result files of every treatment.

Stop Processing

The **Stop Processing** command stops the active process.

The Analysis menu

Sonogram Analysis

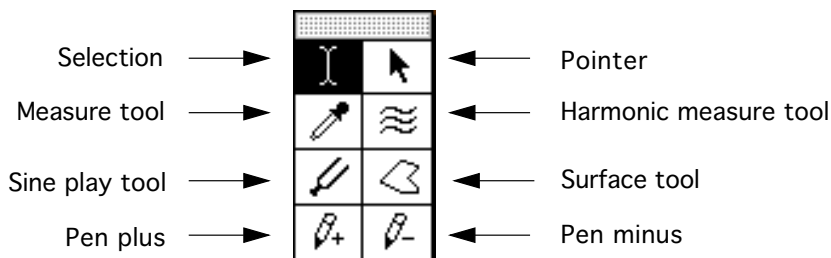
The Sonogram Analysis command performs a spectral analysis of the soundfile. A dialog lets you specify the analysis parameters:

Window size	the window sizes determines the number of sound samples in each analysis.
Fundamental frequency	allows to compute the window size automatically. If you specify the window size, the fundamental frequency that is calculated will be displayed. The window size is set to five times the length of the period that corresponds to the fundamental frequency.
Window step	the window step determines the time interval between two successive analyses.
FFT Size	the FFT size determines the number of points of the analysis.
Analysis window	in the analysis window menu you can specify the window type used in the analysis.
Analysis type	two types of spectral analysis techniques are proposed: fast Fourier transform and linear prediction.

The **Factory Settings** button sets the parameters to the default value.

On launching the analysis, two new panes are added to the window beneath the existing sound pane: the left pane is the sonogram pane, the right pane is the short-term spectrum pane. You can stop the analysis with the **Stop** button in the progress window or by selecting **Stop Sonogram Analysis** in the Analysis menu.

After the analysis, a toolbox is brought on the screen. Here follows a short description of the tools:

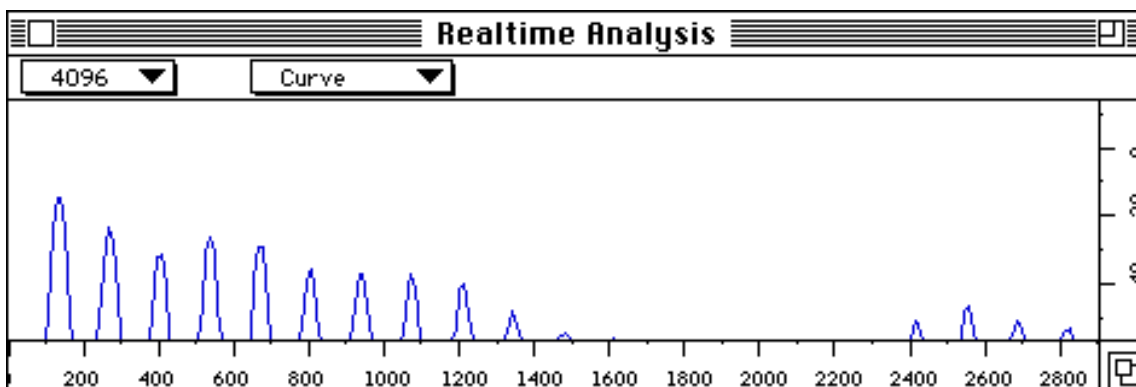


The Selection tool	With the Selection tool (also called linear selection tool) you can select a region in the soundfile.
The Pointer tool	With the pointer tool (also called arrow selection tool) you can select and drag graphical objects drawn on the sonogram. With the pointer tool break-points of a breakpoint function can be moved.
The Measure tool	The measure tool gives you the coordinates of the cursor position in the sonogram.
The Harmonic measure tool	The Harmonic measure tool shows the first 18 harmonic partials of a fundamental frequency in the form of horizontal lines on the screen.

The Sine play tool	The sine play tool plays a sine tone with a frequency and amplitude corresponding to the clicked point in the sonogram. This tool works only on Power Macintosh.
The Surface tool	With the surface tool you can draw surfaces on the sonogram pane (see also the Sonogram Filter). These surfaces specify the regions of the sonogram that will be filtered.
The Pen plus	With the pen plus (also called Pass pencil) you can draw freely on the sonogram to mark the regions that will be passed by the sonogram filter (see Sonogram Filter).
The Pen minus	With the pen minus (also called Cut pencil) you can draw freely on the sonogram to mark the regions that will be eliminated by the sonogram filter (see Sonogram Filter).

Realtime Analysis

This command works only on Power Macintosh. The **Realtime Analysis** command lets you analyze sound in realtime. You must have a microphone or any sound device connected to your computers sound input. The incoming sound will be analyzed (FFT analysis) and the spectrum will be displayed in a new window. You can set the number of points of the analysis in the menu on the left just below the title bar. In the menu on the right, just below the title bar, you can change the way the spectrum is displayed: as a curve, as a series of lines or as a sonogram. The frequency axis is shown below; on the right you find the amplitude axis in case the spectrum is displayed as a curve or as lines. Both axis can be rescaled.



Analysis Settings

The **Analysis Settings** dialog lets you specify the parameters for the FFT analysis and the synthesis:

Window size	the window sizes determines the number of sound samples in each analysis.
Fundamental frequency	allows to compute the window size automatically. If you specify the window size, the corresponding fundamental frequency will be displayed. The

	window size is set to five times the length of the period that corresponds to the fundamental frequency.
Window step	the window step determines the time interval between two analyses. The time interval is measured in number of samples.
FFT Size	the FFT size determines the number of points of the analysis. The number of points in which the spectrum is displayed is always the FFT size divided by 2.
Analysis window	in the analysis window menu you can specify the window type used in the analysis. Three types of windows are proposed: Hanning, Hamming and Blackman.
Synthesis window	in the synthesis window menu you can specify the window type used in the synthesis. Two types of windows are proposed: Hanning and Hamming.

The **Factory Settings** button sets the parameters to the default values.

Realtime Sonogram Scaling

If the **Realtime Sonogram Scaling** command is marked the sonogram will be re-drawn while you re-scale the time or frequency axis. If the command is unmarked the sonogram will temporarily disappear while you rescale. This allows a faster rescaling on slower Macintoshes. To mark the command you click it once; to unmark it you click it a second time.

Edit Color Palette

The **Edit Color Palette** command brings the color palette onto the screen. With the color palette you can change the contrast of the sonogram. You can choose between four types of black and white or color contrasts in the menu in the top of the palette. The minimum and maximum display threshold (corresponding to the minimum and maximum of the grey scale) can be changed with the help of the two triangular markers in the bottom of the palette. The thresholds are measured in dB. The curve in the frame in the centre of the palette defines the dB to grey scale conversion to be applied between the two thresholds. The default curve is linear but a new curve can be drawn with the mouse. The stripe above the frame displays the resulting color conversion. The changes in the color palette effect all the sonogram windows on the desktop.

FFT Analysis

The **FFT Analysis** command launches a short-term spectral analysis using the FFT algorithm. If a part of the sound is selected, only this part will be analyzed; if not, the whole sound will be analyzed. The results are saved as a text file which can then be read by other applications like Max, PatchWork, Excel or a text editor. The parameters for the analysis can be set: the window size or fundamental frequency, the window step, the FFT size, and the analysis window. The results that will be exported and their output format can be defined with the **Output format** button. The

output formats proposed for the FFT Analysis are:

- time (in seconds)
- number (1 plus the FFT size divided by 2)
- freq (instantaneous frequency in Hertz)
- midicents (instantaneous frequency in midicents)
- amp (the amplitude spectrum)
- amp_db (amplitude spectrum in decibels)
- real (the real part of the spectrum)
- imag (the imaginary part of the spectrum)
- phase (the phase spectrum).

In the text file you'll find the time or number in the first line and underneath the requested data in columns separated by a tab.

LPC Analysis

The **LPC Analysis** command launches spectral analysis using linear prediction. It is used to calculate the spectral envelope. If a part of the sound is selected, only this part will be analyzed; if not, the whole sound will be analyzed. The results are saved as a text file which can then be read by other applications like Max, PatchWork, Excel or a text editor. You can define the order of the LPC Analysis as well as the analysis parameters with the **Analysis Parameter** button. The results that will be exported and their output format can be defined with the Output format button. The output formats proposed for the LPC Analysis are:

- time (in seconds)
- number (1 plus the FFT size divided by 2)
- freq (frequency scale)
- midicents (frequency scale in midicents)
- amp (the amplitude of the spectrum envelope)
- amp_db (amplitude of the spectrum envelope in decibels).

In the text file you'll find the time or number in the first line and underneath the requested data in columns separated by a tab.

Peak Detection

The **Peak Detection** command looks for the maxima in the spectra in each consecutive short-term amplitude spectrum. Opening this module gives access to a dialog in which the parameters can be given:

Threshold in dB

the difference in amplitude between a peak at the i'th index and the amplitudes at the neighboring points at indexes i-3, i-2, i+2, i+3

	must exceed the threshold. Peaks that do not satisfy the constraint are neglected. The default value is 25 dB.
Number of Peaks	You can specify how many peaks you want to be noticed. Only the <i>m</i> peaks with the highest amplitude will given as result. By default the 5 highest peaks are given.

With the **Analysis Parameters** button you can set the values of the analysis parameters. If the sonogram is displayed, the results will be shown on the screen as a breakpoint function, if the sonogram is not displayed the results will be registered in a text file. In that case, the formats of the results that can be chosen using the **Output Format** button are:

- time (in seconds)
- number (the number of peaks desired)
- amp (the amplitude of the peak)
- amp_db (the amplitude of the peak in decibel)
- freq (frequency of the peak)
- midicents (frequency of the peak in midicents).

The peaks are ordered according to decreasing amplitude.

Fundamental Estimation

The Fundamental Estimation command looks for the fundamental frequency based on the spectral analysis. In every analysis window an estimation of the fundamental frequency is made. In the dialog window the analysis parameters can be set:

Minimum fundamental	the analysis will not seek the fundamental frequency beneath the threshold specified by the minimum fundamental (the default threshold is 50 Hz).
Maximum fundamental	the analysis will not seek the fundamental frequency above the threshold specified by the maximum fundamental (the default threshold is 2500 Hz).
Maximum harmonic	the analysis will not consider harmonics above the threshold specified by the maximum harmonic (the default threshold is 7500 Hz).
Threshold in dB	specifies a noise level: if the difference in amplitude between a peak and the highest peak is bigger then the threshold, the peak will be neglected.

The ratio Maximum harmonic/Minimum fundamental is limited for reasons of memory storage and calculation time.

If the sonogram is displayed, the result will be shown on the screen as a breakpoint function. If the sonogram is not displayed, the result will be saved in a text file. In that case, the format of the result can be set using the **Output Format** button: **time** (in seconds), **number** (the number of peaks will be 1), **amp** (the amplitude of the fundamental), **amp_db** (the amplitude of the fundamental in decibel), **freq** (the fundamental frequency), **midicents** (the fundamental frequency in midicents). The

parameters for the spectral analysis can be set using the **Analysis Parameter** button.

Masking Effects

The Masking Effects command reduces the amplitude spectrum to a small number of peaks which contribute to the tonal perception. It models the masking effects of the filtering in the internal ear.

In the dialog four different analysis can be chosen:

Weights	the weights measure the spectral importance of the spectral components. The weight is always larger then zero.
Masked amplitudes	the amplitudes of the spectral components after the masking effects have been taken in account. The amplitudes of the significant spectral components have a positive value.
Perceptual pitches	gives an estimation of the "True Spectral Pitch": the pitch of the spectral components as they are perceived by the listener.
All	exports the weights, the masked amplitudes and the perceptual pitches.

Further parameters are:

Threshold in dB	the difference in amplitude between a peak at the i'th index and the amplitudes at the neighboring points at indexes i-3, i-2, i+2, i+3 must exceed the threshold. Peaks that do not satisfy the constraint are neglected. The default value is 25 dB.
Number of Peaks	You can specify how many peaks you want to be registered. Only the m peaks with the highest amplitude will given as result. By default the 5 highest peaks are given.

If the sonogram is displayed, the Perceptual Pitch will be shown on the screen as a breakpoint function, if the sonogram is not displayed the results will be saved in a text file. In that case, the formats of the result can be set using the **Output Format** button: **time** (in seconds), **number** (the number of spectrum peaks), **amp** (the amplitude of the spectrum peaks), **amp_db** (the amplitude of the spectrum peaks in decibel), **freq** (the frequency of the spectrum peaks), **midicents** (the frequency of the spectrum peaks in midicents), and, depending on the requested analysis: **weights**, **masked amplitude** or **perceptual pitch** of the peaks. The parameters for the spectral analysis can be set using the **Analysis Parameter** button.

The Sound menu

Play Sound

The **Play Sound** command plays the complete soundfile of the active window. You can also press the option key and the space bar simultaneously.

Play Selection

The **Play Selection** command only plays the sound that has previously been selected (the selected sound is highlighted in black). Pressing the space bar has the same function.

Get Sound Info

The **Get Sound Info** command displays a window with the information about the soundfile. At the top of the window you will find the name of the file and its format. AudioSculpt recognizes and uses AIFF, AudioSculpt and Sound Designer II formats. The data format is 16 bits integer or floating point in the case of an AudioSculpt soundfile. The information window also shows the sampling rate, the length and the number of samples of the soundfile.

Sound Output

Be sure that your Macintosh is equipped to reproduce sound with the proper quality. The soundfile can be played using the Apple Sound Manager version 3. If your computer is equipped with the Digidesign Card, it can be used to play the soundfile. See the AudioSculpt Newsletter to find out more about playing back soundfiles.

The Windows menu

In the **Window** menu you'll find the list of all the soundfiles that are actually open. The active soundfile is marked. If you choose another soundfile in this menu, its window will come to the front and will be made the active window.

Show Console

The **Show Console** command displays the console window on the screen.

Annex A: Command lines

All the commands of AudioSculpt that are available through the menu bar, can be launched using the command line. The only exceptions are the sonogram analysis and the sonogram filter. Certain processing options are available only through the command line and are not available as through the commands in the menus. This annex gives the equivalent command line of the commands in the menus.

The command line always starts with the identifier "svp", followed by a number of options. An option is preceded by a minus sign "-" and can be followed by a set of parameters and values. The options can be given in any order, though the last parameter should be the name of the output file. Since AudioSculpt can treat two files at once, the options concerning the first sound file are written in capital letters and the those for the second file in small letters. The options are explained in more details in annex B.

Linear prediction analysis (LPC analysis)

```
svp -v -t -S"name soundfile" -Alpc 30 -U -Oa -N4096 -M4000 -Whanning  
-I2000 "name outputfile"
```

Fast Fourier transform (FFT analysis)

```
svp -v -t -S"name soundfile" -Afft -U -Oa -N4096 -M4000 -Whanning -  
I2000 "name outputfile"
```

Peak detection

```
svp -v -t -S"name soundfile" -Apic 0.0 n5 -U -Oa -N4096 -M4000 -Whan-  
ning -I2000 "name outputfile"
```

Fundamental Estimation

```
svp -v -t -S"name soundfile" -Af0 fm50, fM2500 F7500, sn50 -U -Oa -  
N4096 -M4000 -Whanning -I2000 "name outputfile"
```

Masking effects

```
svp -v -t -S"name soundfile" -Amask weight 15.0 n10 -U -Oa -N4096 -M4000 -Whanning -I2000 "name outputfile"
```

Band filter

```
svp -v -t -S"name soundfile" -Fbande "parameterfile" -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Breakpoint filter

```
svp -v -t -S"name soundfile" -Fbreakpt "parameterfile" -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Formant filter, interpolation on the parameters

```
svp -v -t -S"name soundfile" -Ffof "parameterfile" -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Formant filter, interpolation on the filter responses

```
svp -v -t -S"name soundfile" -Ffifof "parameterfile" -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Surface filter

```
svp -v -t -S"name soundfile" -Fsurface "parameterfile" -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Cross synthesis source filter

```
svp -v -t -S"filtered sound" -s"filtering sound" -Gmul -A -Z -Jhamming -N4096 -M4000 -Whanning -I2000 -alpc 30 -whamming -m1024 -i128 "name outputfile"
```

General cross synthesis

```
svp -v -t -S"filtered sound" -s"filtering sound" -Gcross -X1.0 -x1.0 -Y1.0 -y1.0 -q0.0 -A -Z -Jhamming -N4096 -M4000 -Whanning -I2000 -a -whamming -m1024 -i128 "name outputfile"
```

Sample rate conversion

```
svp -v -t -S"name soundfile" -H22050 -resWhanning -resM21 "name outputfile"
```

Time compression/expansion

```
svp -v -t -S"name soundfile" -D2 -A -Z -Jhamming -N4096 -M4000 -Whanning -I250 "name outputfile"
```

Transposition without time correction

```
svp -v -t -S"name soundfile" -trans 1200 -resWhanning -resM21 "name outputfile"
```

Transposition with time correction

```
svp -v -t -S"name soundfile" -D -trans 1200 -A -Z -Jhamming -N4096 -M4000 -Whanning -I2000 -resWhanning -resM21 "name outputfile"
```

Annex B: List of options of the command line

A value given between brackets shows the default value of a parameter. Additional examples of command lines can be found in annex A.

-A<analysis type> or -a<analysis type>

Specifies the analysis type. THE Following types can be used:

- **fft**: Fast Fourier Transform

Example: -Afft

- **lpc<order>** linear prediction analysis, with
<order> the number of poles [30].

Example -Alpc50

- **pic<threshold><nnumber>**: peak detection, with
<threshold> threshold is given in dB, and
<number> the number of peaks to be detected.

Example: -Apic 3 n20

- **mask<list><threshold><nnumber>**: Terhardt's masking algorithm, with
<list> of parameters: masked amplitudes (amp), perceived frequencies, weights (weight), or all of the previous (all),
<threshold>: threshold beyond which peaks are sought [25dB], and
<number>: the number of peaks to be detected.

Example: -Amask amp, weight 10 n20

- **f0<fmmin><fMmax><FMAX><snthreshold>**: fundamental estimation, with
<min> minimum fundamental frequency [50],

<max> maximum fundamental frequency [2500],
<MAX> maximum harmonic frequency [7500], and
<threshold> threshold is given in dB [50].

Example: -Af0 fm100, fM1000, F2500, sn60

-B<start> or -b<start>

Specifies the start time of the analysis within the sound file, with

<start> given as sample number (integer) or as time in seconds (float). The start time can be negative to add silence at the beginning.

Example: -B44100, -B1.4

-D<value> or <file name>

Constant or time variable compression/expansion, with

<value> the compression/expansion factor, and

<file name> the name of the file describing the compression/expansion factor.

Example: -D2.0 (expansion with factor 2 = sound file will be twice as long)

In case a transposition with time correction is wanted, this option should be included in the command line. No value for the compression/expansion factor should be given since AudioSculpt will calculate it automatically.

-E<end> or -e<end>

Specifies the end time of the analysis within the sound file, with

<end> given as sample number (integer) or as time in seconds (float). The end time can be bigger than the length of the sound file to add silence at the end.

Example: -B88200, -B5.0

-F<filter type><file name> or -f<filter type><file name>

Determines the filter type. The following types are available:

- breakpt: breakpoint filter
 - bande: band filter
 - surface: surface filter
 - fifof: formant filter, interpolation of the filter response
 - fof: formant filter, interpolation of the filter parameters
- <file name> contains the temporal evolution of the parameters.

Example: -Fbande ParaFile

**-G<cross synthesis type><file name> or <cross synthesis type><-Xvalue>
<-xvalue><-Yvalue><-yvalue><-qvalue>**

The two types of cross synthesis that can be used are:

- cross: general cross synthesis
- mul: source/filter cross synthesis

The parameters can be given in a file:

<file name> contains the parameters for a variable cross synthesis, or directly in the command line (constant cross synthesis):

<X> factor of the amplitude spectrum of the first sound [1.0]

<x> factor of the amplitude spectrum of the second sound [1.0]

<Y> factor of the phase spectrum of the first sound [1.0]

<y> factor of the phase spectrum of the second sound [1.0]

<q> factor of comultiplication [0.0]

Example: -Gcross CrossPara

-g<gain><file name> or <tremolo><mode><file name>

Performs an amplitude modulation. The modulation envelopes are described in a file containing the time/amplitude information. The following types are available:

- sinus: sinus wave modulation
- carre: square wave modulation
- triangle: triangular wave modulation
- scie: sawtooth wave modulation

Examples: -ggain MyFile, -gtremolo sinus MyFile

-H<sample rate>

Determines the new sampling rate of the sound file. If this option is used, do not use the options -A and -Z.

Example: -H32000

-h

Displays the help file containing a description of the options.

-l<step> or <method><file name>, or -i<step> or <method><file name>

Specifies analysis step in samples. The default values is 1/8 of the window size.

<method> is the method for reading the parameter <file name>. Three methods are possible:

- sync: ordered pairs in parameter file (time, fundamental frequency)
- dep: ordered pairs in parameter file (time, step value)
- pos: sample position in the sound file is given for each window

Examples: -l256, -lsync StepFile

-J<window type>

Specifies synthesis window type. There are two possible windows:

- hamming (default)
- hanning

-M<window size> or <method><file name>, or -m<window size> or <method><file name>

Specifies analysis window size, with

<window size>: expressed in samples [1024], or

<method>: the method for reading the parameter file <file name>. The following modes are possible:

- sync: ordered pairs in parameter file (time, fundamental frequency)
- dep: ordered pairs in parameter file (time, window size)

Examples: -M2000, -Mdep MyFile

-N<FFT size> or -n<FFT size>

Specifies FFT size in samples [1024]. The FFT size must be greater than window size (see -M).

Examples: -N4096, -N8192

-nn-

Do not normalize output sound file (default is normalize)

-ns

Spectral normalization for analysis output (default is no normalization)

-O<mode>:<list> or <mode><file name>

Specifies the output format, with

<mode> is the output type. Possible modes types are:

- a for ASCII
- b for binary (default)

<list> is a list of strings separated by ',' (number, time, amplitude, frequency, phase, etc.)

<file> is a text file where output format is specified

Examples: -Oa:number,amplitude,phase

-Oa foo.format

-R<sample rate> or -r<sample rate>

Specifies sampling rate. Used when performing resynthesis on analysis data file.

-resM<value>

Specifies resampling window size, with

<value>: the resampling convolution window size in samples [31]

-resW<value>

Specifies resampling window type, with

<name>: the resampling window type [hamming]

-S<file name> or -s<file name>

Specifies input file name, read from the default sound folder.

Example: -Sflute

-T

Display the current position of the synthesis window during the calculation.

-t

Display the current position of the analysis window during the calculation.

-trans<cents>

Transposes by the given number of cents.

Time correction is performed if used in conjunction with the flags -A -Z -D

Example: -trans 1200 (will transpose one octave up)

-v

Prints out details about the AudioSculpt patch structure and parameters.

-W<window type> or -w<window type>

Specifies analysis window type. There are three possible windows:

- hamming (default)
- hanning
- blackman

Example: -Whanning

-Z

Performs a resynthesis (inverse fft and overlap/add).

Index

A

AIFF 14, 89
amp 25
amp_db 25
Amplitude spectrum 51, 64
Analysis menu 82
Analysis parameters 23
Analysis settings 84
Analysis type 83
Analysis window 83, 85
Apple 2
Audiomedia II 13

B

Band filtering 55, 79
Battier M. 2
Blackman 35, 38, 85
Breakpoint filtering 56, 79

C

Carrive J. 2, 11
Carrive, J. 7
Clear 76
Close 75
Command line 82, 91, 94
Constant transposition 23
Copy 76
Cross synthesis
 generalized 62
 source-filter 62, 63, 81
Cross synthesis source-filter 81
Cross-synthesis
 generalized 81
Cut 76
Cut pencil 84

D

Depalle Ph. 2, 11, 36
Depalle, Ph. 7
Desired sampling rate 73
Digidesign 13, 89
Doval B. 2

E

Eckel G. 2, 11
Eckel, G. 7
Edit color palette 85
Edit menu 76
Excel 39, 85, 86

F

Factory settings 38, 83
Fast Fourier transform 33
FFT 8, 33
 analysis 23, 24, 85
 size 26, 33, 35, 83, 85
Fichier 15
File 15, 74
Filter
 band 51, 52, 55, 79
 breakpoint 51, 52, 56, 79
 fof 59
 formant 52, 58, 80
 pass-band 51
 sonogram 61, 81
 stop-band 51
 surface 51, 52, 60, 80
Filter response 28
Fof 59
Fondamental 8
Formant filter 28, 58
Formant filtering 80
freq 25
Fundamental Estimation 8, 87
Fundamental estimation 43
Fundamental frequency 83, 84

G

Garcia G. 2
Generalized cross synthesis 81
Get sound info 89

H

Hamming 35, 38, 85
Hanning 35, 38, 85
Harmonic measure tool 39, 83

I

imag 25

L

Linear prediction (LPC) 36
Linear selection tool 83
LPC 8, 33, 86

M

Mary M. 14
Masked amplitudes 45, 46, 88
Masking Effects 45, 88
Masking effects 45
Max 39, 85, 86

Maximum fundamental 44, 87
Maximum harmonic 44, 87
Measure tool 38, 83
midicents 25
Minimum fundamental 44, 87

N

New 74
New Text Document 82
number 25
Number of peaks 87, 88

O

Open 15, 75
Option
-A 94
-B 95
-D 95
-E 95
-F 95
-G 96
-g 96
-H 96
-h 96
-l 97
-J 97
-M 97
-N 97
-nn 97
-ns 97
-O 97
-R 98
-resM 98
-resW 98
-S 98
-T 98
-t 98
-trans 98
-v 98
-W 98
-Z 99
Options 94
Output format 24
Ouvrir 15

P

Page setup 75
Pass pencil 54, 84
Paste 76
PatchWork 39, 85, 86
Pauset B. 2
PCI 13
Peak Detection 8, 42
Peak detection 45, 86
Pen minus 84
Pen plus 84
Perceptual pitch 45, 46, 88
Phase 25

Phase spectrum 51, 64
Pics 8
Play selection 89
Play sound 17, 89
Plus pen 54
Pointer tool 44, 83
Poirot G. 2
Poirot, G. 7
Poles 36, 38, 63
Power Macintosh 19
Preferences 77
Print... 75
Process and save 22, 23
Process in place 22, 23
Processing menu 21, 78
ProTools 13

Q

Quit 75

R

real 25
Realtime analysis 84
Realtime sonogram scaling 85
Resampling parameters 23, 73
Rodet X. 2
Rogers C. 2, 11
Rogers, C. 7
Run Script 82

S

Sample rate conversion 73, 78
Save 75
Save As... 75
Select All 76
Selection tool 83
Serra M.-H. 2, 11
Serra, M.-H. 7
Set Default Folders 75
Set default folders 13
Short-term spectrum 32
Show Clipboard 77
Show console 89
Sine play tool 39, 84
Sonogram 37, 38, 46, 61
Sonogram analysis 17, 82, 83
Sonogram filter 21, 53, 61
Sonogram filtering 81
Sound Designer II 14, 89
Sound folder 13
Sound Manager 13, 89
Sound menu 88

Sound output 89
Spectral envelope 36, 59
Spectrogram filter 82
Stop processing 82
Stop sonogram analysis 83

Surface filtering 60, 80
Surface tool 54, 84
SVP 7
Synthesis window 85

T

Terhardt 8
Terhardt E. 2, 45
Text file 37, 39
Threshold in dB 44, 86, 87, 88
time 25
Time scale modification 78
Transfer function 79
Transposition 22, 78
True Spectral Pitch 88

U

Undo 23, 76
UNIX 82
Use Command Line 82

V

Vocodeur de phase 8

W

Weights 45, 46, 88
Window size 31, 62, 83, 84
Window step 32, 83, 85