A New Software Package for Spectral Investigation and Analysis/Synthesis Using FFT and Sinusoidal Modelling Techniques

Chris Scallan
Music Department, La Trobe University
Bundoora, Australia
Email: cscallan@klag.latrobe.edu.au

Thomas Stainsby
Music Department, La Trobe University
Bundoora, Australia
Email: stainsby@klag.latrobe.edu.au

Abstract

AnnaLies 4.0 is the latest version of an original software package written by the authors for the investigation and manipulation of sounds in the frequency domain. Sounds can be analysed using both Fourier and sinusoidal-modelling techniques, with the resultant data displayed in a variety of formats, in both two and three dimensions. This data can also be used in the synthesis of sounds with a variety of modifications. This paper presents a quick summary of the theoretical background, a stepwise demonstration of the program and a brief discussion of its further applications, including work in audio signal separation.

1 AnnaLies 4.0

In its present form, AnnaLies 4.0 has already been used as a tool for spectral investigation, as well as providing a unified environment for the development of further analysis algorithms, such as those required for Thomas Stainsby’s current research in mixed signal decomposition. Co-authored by Chris Scallan and Thomas Stainsby, this program began as a rewrite of AnnaLies, which was written in 1991 by David Hirzt and Thomas Stainsby, itself incorporating code written by Graeme Gerrard (1990). The current version of the program has gone well beyond the original FFT analysis version, and offers a variety of display options, as well as incorporating MQ analysis and resynthesis.

2 Theoretical Background

AnnaLies 4.0 initially performs an FFT analysis of a sound, with the ability to perform further analysis procedures, MQ Analysis, on the FFT output data. The following is a brief outline of the theoretical foundations of these procedures.

2.1 FFT Analysis

The term FFT analysis refers to a specific digital implementation of the more generalised conceptual procedure of Fourier Analysis. This type of analysis represents a complex waveform in terms of its frequency spectrum components, consisting of a finite series of sinusoids with frequencies at whole-number multiples of the fundamental frequency. Fourier’s theorem states that any waveform can be represented as such a manner. The FFT algorithm is a specifically optimized form of the Discrete Fourier Transform (DFT), which represents a discrete signal in terms of a discrete frequency spectrum. The FFT achieves a greatly reduced calculation time by virtue of the symmetry inherent in the output data. The computation is significantly reduced further when dealing exclusively with real signals, as is the case in audio.

A detailed discussion of the mathematics involved in the FFT algorithm would be inappropriate here, so for our discussion it is only necessary to understand that the output of the FFT analysis procedure is a representation of a sound’s frequency content as a quantised frequency spectrum, with frequency channels spaced at regular intervals, typically in the range of 25 to 100 Hz apart. The frequency resolution is determined by the length of the transformation window, and is equal to the sampling frequency divided by the transform length. The underlying trade-off in the choice of values for the analysis parameters is that an improved frequency resolution requires a decreased temporal resolution. For further information on FFT analysis, an excellent discussion can be found in Moore (1990: 61 – 111).

The analysis of sound can be taken one step beyond representation in terms of a quantised spectrum, to represent sounds in a manner closer to the way our ears naturally interpret them, namely as a set of individual frequency components. The next analysis procedure implemented in the AnnaLies 4.0 program, termed MQ analysis, produces such a model of the sound as a set of individual sinusoidal components.

2.2 MQ Analysis

The MQ analysis procedure was first described by McAuley and Quaylen (1986), after whom it has been named (see also Quaylen and McAuley 1986). This procedure constructs an additive sinusoidal model of...
the sound where frequency components are represented as independent time-varying sinusoidal partials.

The frequencies and amplitudes of these partials are inferred directly from the time varying spectra represented by the FFT procedure. We can think of any given peak in one frame of an FFT analysis as representing an underlying sinusoidal component at that point in the spectrum at that point in time. A series of peaks in successive FFT snapshots charts the progress of that component over time. We can determine the exact frequency and amplitude of each spectral peak using parabolic interpolation, and smooth between the different values at each point in time, to give us a time-varying representation of that partial's behaviour [Smith and Serra 1987]. By performing the above interpolation for each significant peak in the frequency spectrum, we can assemble an additive model which represents the behaviour of all the significant frequency components contained in the signal. A component is judged to be significant if its amplitude is above a specified threshold. The MQ analysis thus yields a more intuitive temporal representation of a sound's frequency content.

Anal.Lies.0 also allows the user to verify the results of the MQ analysis aurally, by providing a range of resynthesis options from the MQ data. These include the resynthesis of grouped or selected partials using an additive synthesis procedure.

3 Operation of AnnaLies.4.0

The following is a step-by-step outline of Anal.Lies.0, which runs on the Macintosh computer and has as its audio input sound files prepared using Digidesign’s Sound Designer II program. An attempt to adapt Anal.Lies.0 which distinguishes it from other spectral analysis packages already available is its ability to deal with sound files of virtually any length, by employing its own FFT file data format, as opposed to the conventional method of reading and displaying only one section of a longer sound file at a time. This feature offers the potential of larger scale spectral analysis.

3.1 Demonstration of AnnaLies.4.0

1) Firstly a sound file must be created using Sound Designer II. For this demonstration, we have recorded a flute and 'cello diad. This sound file was prepared for use in Thomas Stainsby's current research into mixed signal decomposition.

2) Upon opening the sound file within Anal.Lies.0, the Analysis Parameters Dialog Box appears, which is used to set the desired analysis parameters. The Window Length parameter determines the frequency resolution provided by the analysis, the user specifying the number of samples contained in the analysis window. The Hop Size parameter determines the spacing in time between successive FFT ‘snapshots’. An appropriate value here can help compensate for the temporal resolution lost by selecting a large window length. The Window Type parameter allows the user to determine the type of mathematical weighting function which will be applied to the samples to be analysed. This is necessary to help compensate for inaccuracies introduced by the FFT analysis procedure as a by-product of its averaging of values over time. As no ideal weighting function exists, the user is given the option to choose the one most appropriate for their needs. The Start Sample and End Sample parameters allow the user to specify which portion of the sound file they wish to analyse.

3) Immediately prior to initiating the analysis, the user creates an “FFT” file, to which the output data will be written as pairs of amplitude and frequency values for each frequency channel within each FFT analysis frame.

4) Once the analysis is complete, the user can view the analysis data by firstly opening the FFT file, and then selecting the preferred mode of display. The various options offered include both 2 and 3 dimensional FFT and 2 dimensional MQ display. The user has the option of either linear or logarithmic amplitude and frequency scaling, and a choice of four perspectives for the 3D FFT display.

Figures 1, 2 and 3 show three different displays of the flute and 'cello diad.

Figure 1: 3D FFT plot of a flute and 'cello diad.
4 Conclusion and further applications for AnnaLies 4.0

In its present form, AnnaLies 4.0 is probably the most versatile Macintosh package currently available for viewing spectral representations of sounds and graphically editing their spectral content. Future development lies in the area of linking it to more methods of signal processing and synthesis, in which capacity it already provides a platform for the implementation of the signal separation strategies being developed by Thomas Stainsbry.

In the latter research, the approach taken by Robert Maher (1989,1990) has served as a starting place, as he also investigated separating simultaneous sounds represented by a sinusoidal model derived from MQ analysis. Once a mixed signal has been represented in this form, various intelligent decision-making strategies could be employed to determine the source of each spectral component, which would then allow each sound source to be reconstructed independently. T.S.'s latest research is associated with applying more intelligent decision-making algorithms to this problem of partial assignment.

References

Gerrard, Graeme. 1990. "flats\&n", STFT analysis routine contained in AnnaLies project.


Sound Designer II, software package for Macintosh by Digidesign.