Audiophile's Analyzer 2.0
User's Guide

Contents

Introduction

The Audiophile's Analyzer is designed to analyze audio files; and to provide tools to allow the user to investigate the various techniques available to analyze audio files. The metadata will be extracted and displayed. The audio waveform and FFT will be displayed. Spectrograms will be displayed.

For music files a set of tools will allow the user to transcribe the music to musical score, and MIDI format.

The Audiophile's Analyzer 2.0 has all the features of 1.3 and adds spectrograms from multi-rate FFTs and Constant-Q transforms. Version 2.0 includes AI functionality allowing the user to configure, train and use Convolutional Neural Networks to transcribe from spectrograms.

The Audiophile's Analyzer is designed to work on a desktop computer running Windows 11. It will not be supported on Windows XP, Linux, or iOS.

The Audiophile's Analyzer is a product of Sound Analysis.

Installation

Before installing the Audiophile's Analyzer; make sure that you have version 4.8.1 (or later) of the .NET framework installed. If not, you can get it from Microsoft at:

[Download .NET Framework 4.8.1 | Free official downloads (microsoft.com)](#)

The Audiophile's Analyzer is available from the Microsoft Store (search for Audiophile's Analyzer).

Alternatively, the installation files are delivered as a zip file available from the Sound Analysis web site http://sound-analysis.com/home-4/audiophiles-analyzer/ . Extract the file to any location on your computer, and then run setup.exe. If you have a previous version of the Audiophile's Analyzer installed, uninstall that first.

Opening a file

Most audio and video file formats are supported. Files other than .wav files will be converted to .wav and located in the Music\ConvertedWavFiles directory.
MIDI files are supported.

The Open File dialog can be invoked from the menu bar File -> Open,
or audio files can be dropped onto the Audiophile's Analyzer,
or right click on the file, use Open with and select the Audiophile's Analyzer,
or from the command line type SoundAnalysis.AudiophilesAnalyzer.exe followed by the path to the file to open.
The metadata will be displayed.

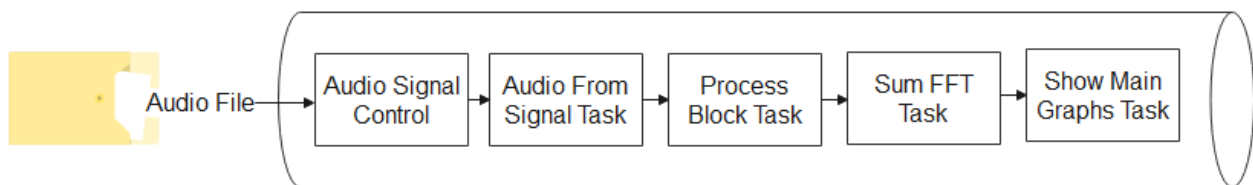### The Data Acquisition Workflow

Figure 1 : Data Acquisition Workflow

The Data Acquisition Workflow is initiated by the file read operation. Each frame / block (8k samples) is read by the Audio Signal Control and passed to the pipeline where the raw signal is extracted, the channels separated, the single block FFT is performed and summed. Finally, the Show Main Graphs tasks will call the Audio signal control to display the envelop. Signal and FFT graphs.
After the main graphs are completed, the multi-rate FFTs will be calculated, a progress dialog will be displayed as this can take a while. Once this is completed the Details, Analysis, and Spectrogram tabs will be enabled. The multi-rate FFTs are limited to the first minute of audio, much more will become unwieldy in memory usage and time to calculate.

Metadata

The file metadata will be displayed at the top of the left-hand panel. The metadata displayed will depend on the metadata contained within the file.

FFT Expanders

Multiple FFTs are performed after the file is loaded. The larger FFTs are so that low notes can be transcribed, the smaller are to improve the timing of the beats. The expanders on the left will describe each FFT size including which octaves can be supported. This information is calculated for each file based on the sample rate and block size.

Tabs Overview

Signals Tab

The envelope, signal, and FFT of the file will be displayed in the Signals tab. The signal may be played back using the play button.
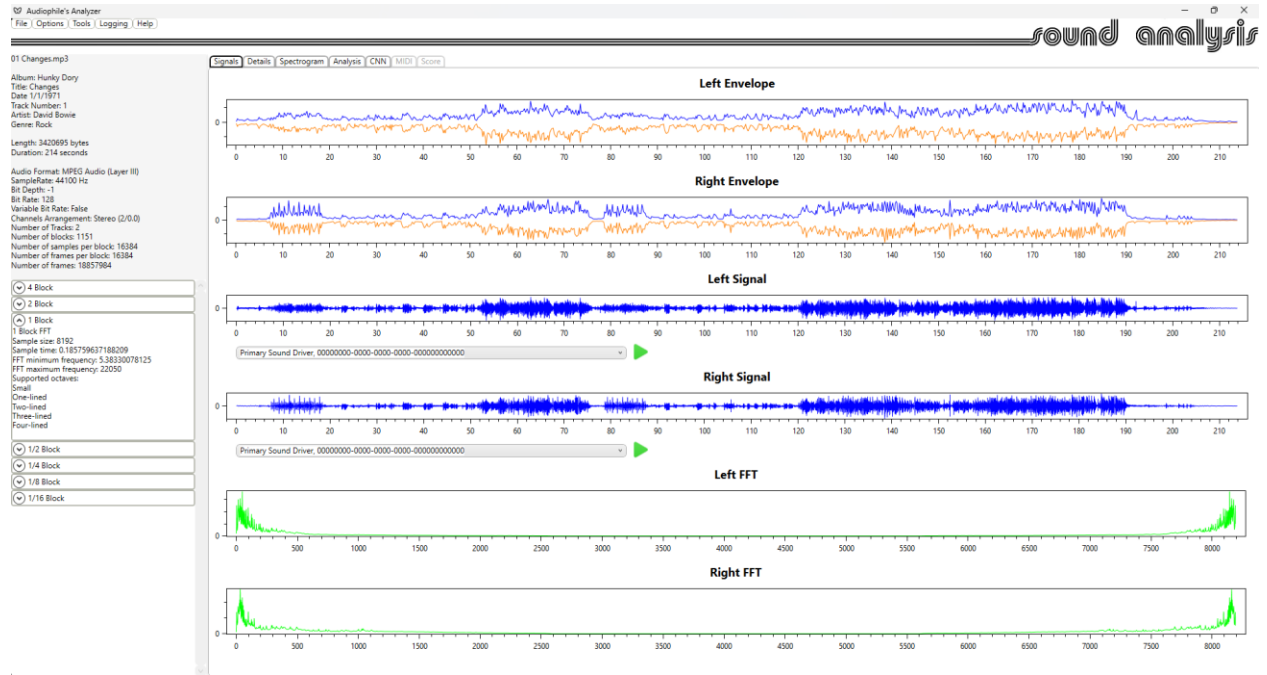
Figure 2 : Signals Tab

The FFT displayed here is the sum of FFTs for every block in the file. This will be used to calculate the file wide default values used to initialize the Analysis tab; for example, the octave range in the Octave Mapping expander, and the Key Signature in the Score Options group.

Details Tab

The details tab will show the signal and spectrum for each block. The slider will allow the user to change the block shown.

Figure 3 : Details Tab

Spectrogram Tab

The spectrograms will provide a visual representation of the audio, and provide for transcription either directly using the Constant Q spectrogram or with a trained Convolutional Neural Network (CNN) using any of the spectrograms (the provided CNN models use the multi-rate FFT with monochrome selected).
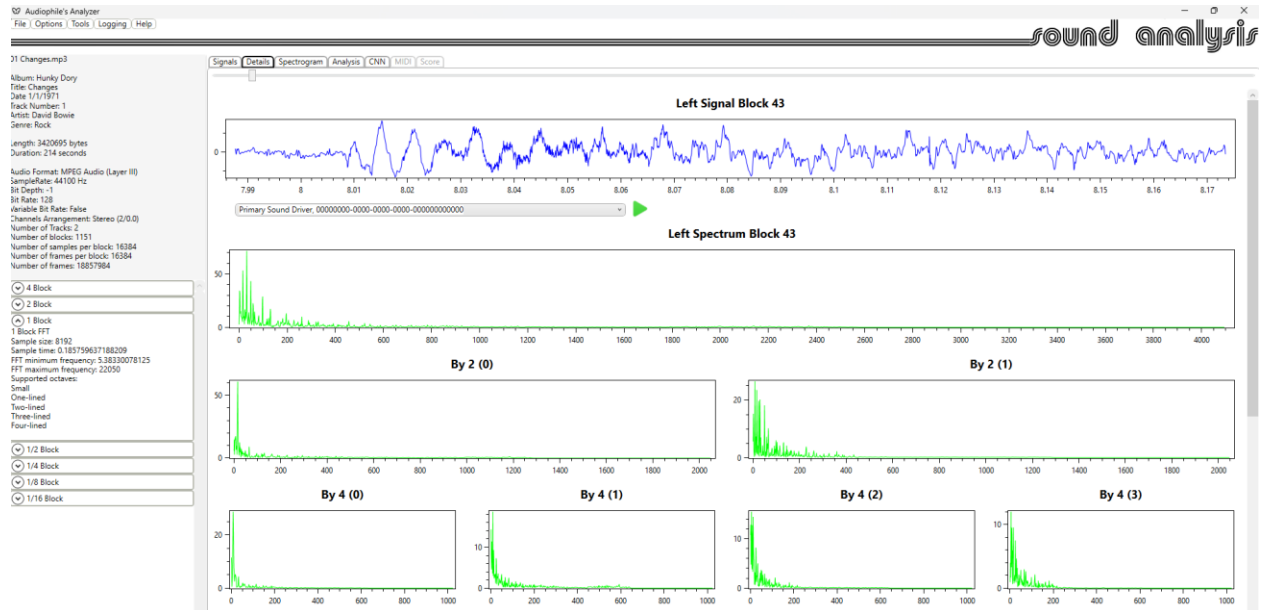
Each type of spectrogram may be displayed as pseudo-color or monochrome.
The user can zoom in and change the FFT size dynamically.
The gain may be varied dynamically.

The width of the spectrogram will depend on the length of the audio file used for input.
The height will depend on the spectrogram type selected:
Single FFT 360 pixels.
Multi-rate FFT 161 pixels.
Mapped notes 60 pixels.
Constant Q 84 pixels.

Single FFT
With FFT selected the Spectrogram tab will show a spectrogram based on any of the FFT sizes available. The horizontal axis is time, the vertical axis is frequency.

The user can zoom in and change the FFT size dynamically. The focus on low frequencies slider will allow the user to trim off the high frequencies, which often have no useful information.
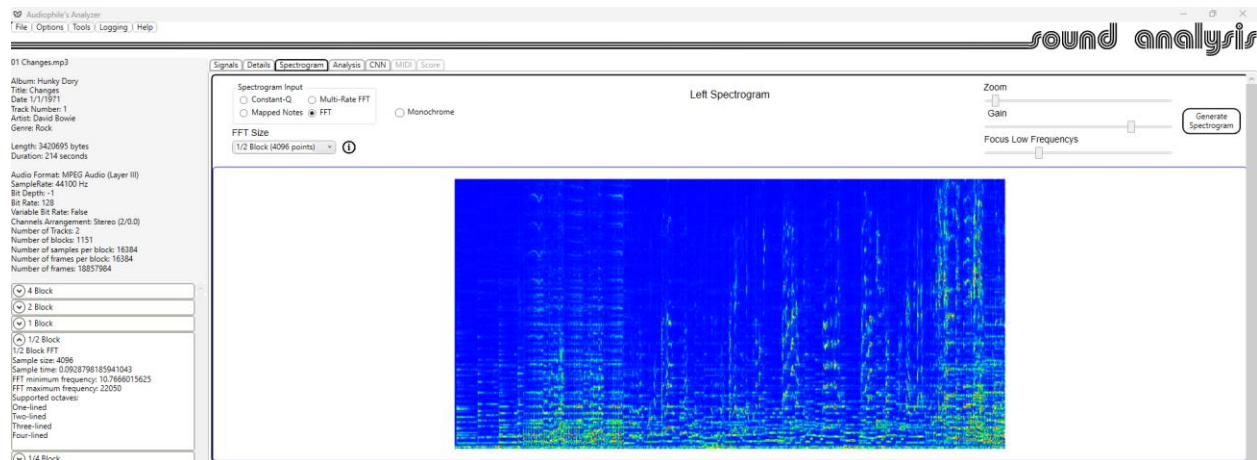
Figure 4 : Spectrogram Tab – FFT selected

Multi-rate FFT

With Multi-rate FFT selected, the Spectrogram Tab will show a spectrogram based on all available FFTs with each FFT being used to cover a single octave.

Each FFT will contribute buckets 24-46 to cover every note in an octave from C to B.

The 4 block FFT will cover octave 1 (Contra)

The 1/16 block FFT will cover octave 7 (Four-lined)

There will be one time-slice for each 1/16 block FFT sample, other FFT contributions will be stretched over multiple time slices.
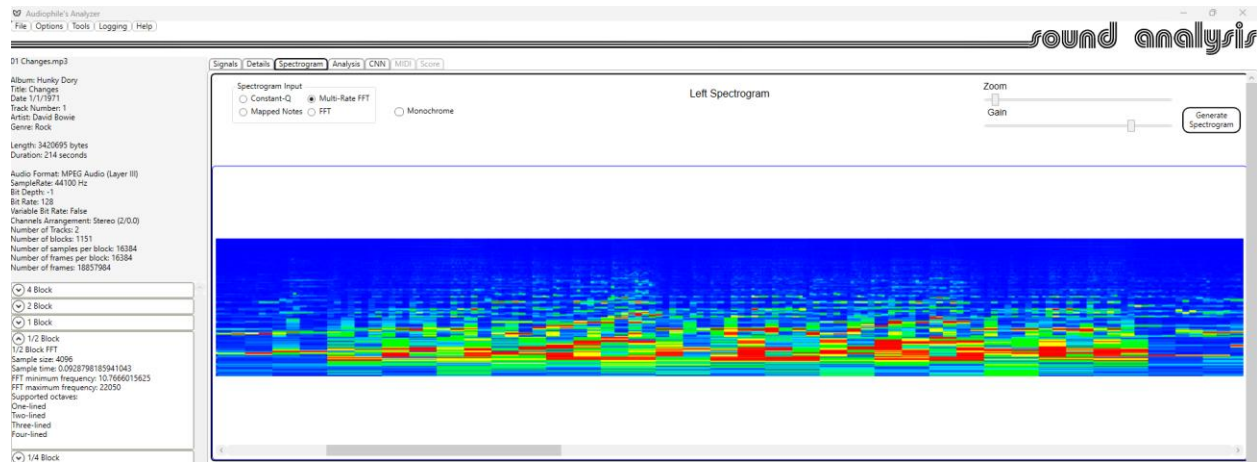
Figure 5 : Spectrogram with Muti-rate FFT selected

Mapped Notes

Selecting the Mapped Notes radio button will display a spectrogram from the mapped notes of the first three octaves allowed for the FFT size selected.

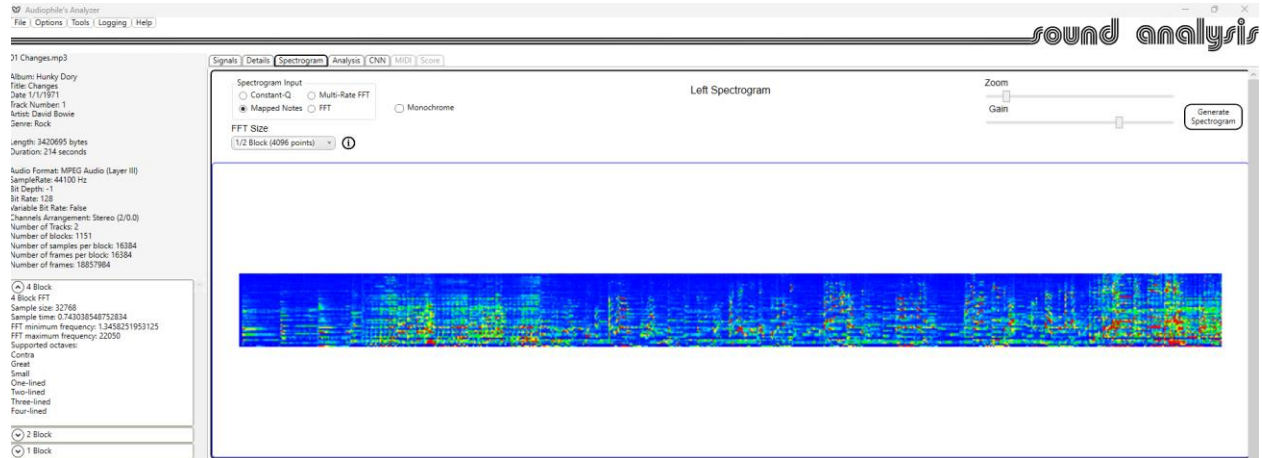**Figure 6 : Spectrogram - Mapped Notes selected**

Constant Q

Selecting the Constant Q radio button will display a spectrogram based on the Constant Q transform. This has 84 buckets one for each note over 7 octaves. This transform is generated from the Multi-rate FFT by apply a hamming window to the FFT results for each note.
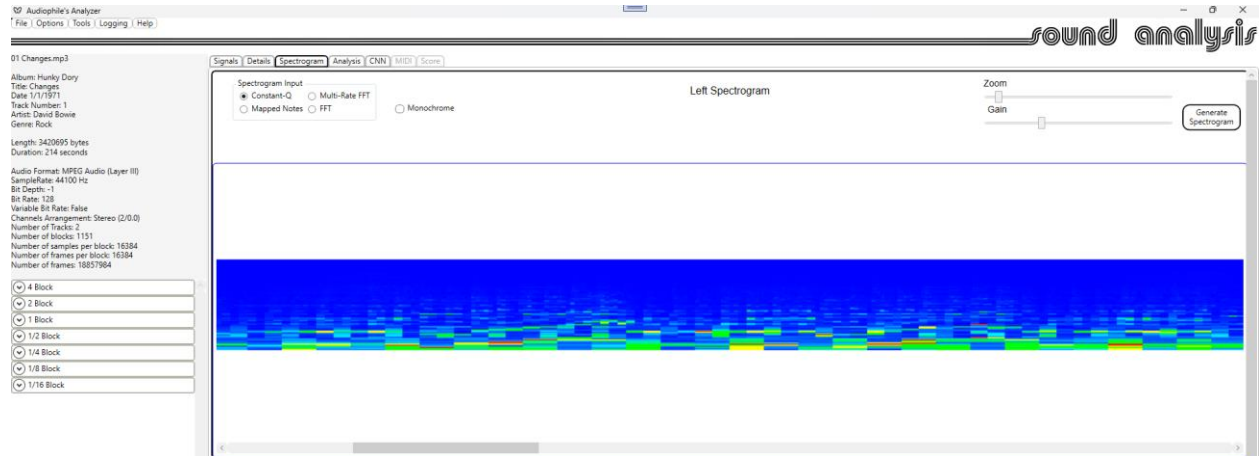

Figure 7:Spectrogram - Constant Q selected

Analysis Tab

The analysis tab will show how the FFT results will be mapped to notes, and allow the user to track the beats from the FFT results.

Figure 8 : Analysis Tab - Octave Mapping

The Lower and Upper octaves will be selected based on the frequencies detected. These may be changed by the user; the results of the change will be displayed.

The FFT size will be selected to allow the best time resolution for the Lower Octave. This will change dynamically, but may also be changed by the user.

The mapping option will select the way that notes are mapped to frequency buckets.

Most Tolerant will map every bucket to the closest note. For the highest notes many buckets will be mapped to each note. For the lowest notes it is possible that some notes will have no buckets mapped.

More Tolerant will apply a Hamming Window function to the Most Tolerant results during transcription.

Tolerant will limit the number of buckets mapped to a note to 3, the exact match and the bucket each side.

Precise will map every note to the closest bucket. All notes will have exactly one bucket mapped. For the lowest notes it is possible that more than one note will share a bucket.

Accepting the default octave range will avoid mapping errors, this is recommended.

Mapping frequency buckets to notes is a reasonably exact science. Accepting the defaults is almost always the best approach. For this reason, the Octave mapping expander is collapsed by default.
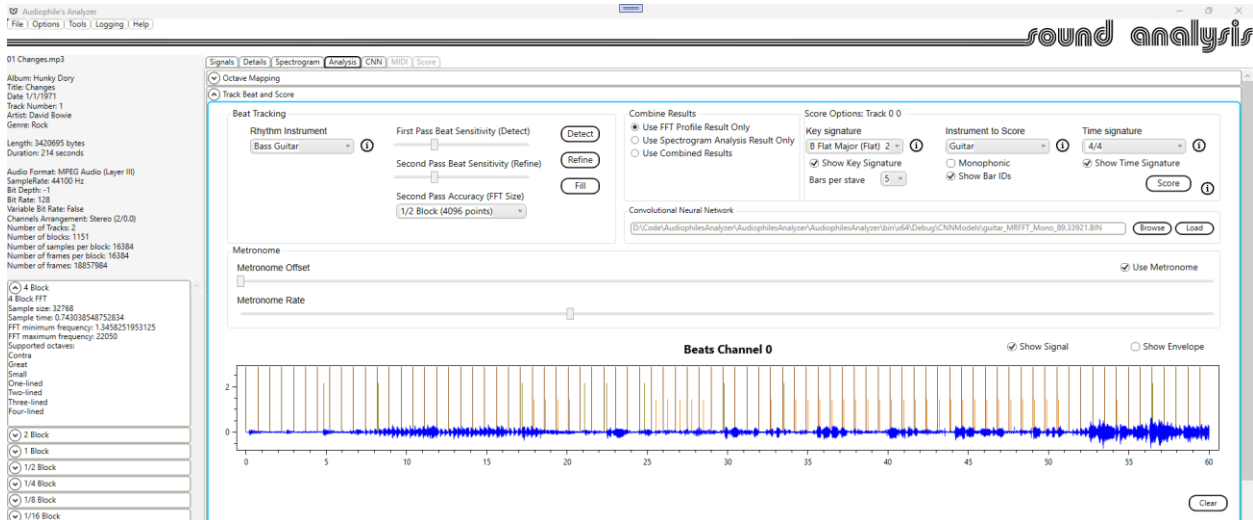
Figure 9 : Analysis Tab - Track Beat and Score

Finding the beats in a piece of music is more of a black art. For this reason, the track beat and score expander is expanded by default. The standard technique for identifying a beat is to look for increases in energy.

The Detect button will use this technique to identify beats using the FFT size selected in the Octave mapping expander.

The Refine button will use a smaller FFT size (selected by the drop down) to locate the beats more precisely. This will default to half the size of the default FFT size in the Octave mapping drop down, but maybe changed by the user.

The sensitivity of these two beat detectors may be adjusted.

The Fill button will try to deduce the tempo and fill in any missing beats based on the beats detected by the Refine button. It will also set the metronome sliders. This step is optional.

Checking the Use Metronome checkbox will show the metronome beats and instruct the Audiophile's Analyzer to use the metronome beats to score. At this stage the metronome sliders can be used to align the metronome beats with the detected and deduced beats.

The metronome is optional and may be used with or without the beat tracker.

The Show Beat Sampling Range will show the signal that will be sampled for each beat. This range will start with beat; it's width will depend on the FFT size selected in the octave mapping expander.

The Show Signal and Show Envelope check boxes will superimpose these signals onto the beats graph for reference.

This step is, by design and necessity, interactive; it may be performed in a few different ways.

Beat Tracking only
1) Select the rhythm instrument.
2) Click Detect (adjust first pass sensitivity if necessary)
3) Click Refine (adjust second pass sensitivity if necessary)
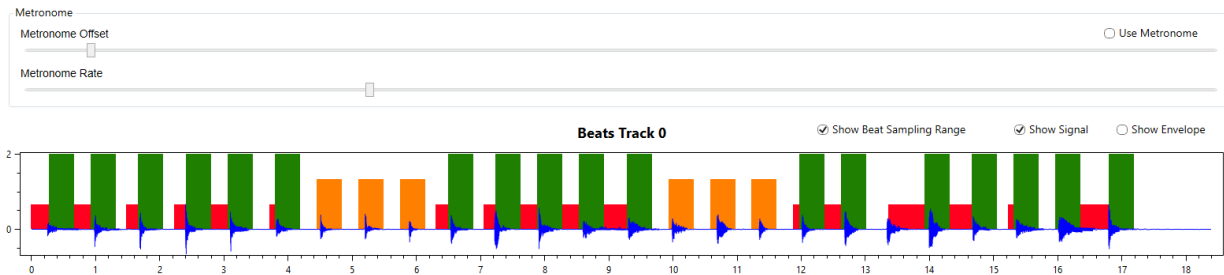4) Click Fill (optional)
5) Set Score Options and click Score.

Figure 10 : Beat detection, beat tracking only

Metronome only
1) Select Use Metronome.
2) Select Show Signal and / or Show Envelope for reference.
3) Use the metronome sliders to adjust the displayed metronome beats.
4) Set Score Options and click Score.

Figure 11 : Beat tracking, metronome only

Beat Tracking and Metronome
1) Select the rhythm instrument.
2) Click Detect (adjust first pass sensitivity if necessary)
3) Click Refine (adjust second pass sensitivity if necessary)
4) Click Fill (required)
5) Select Use Metronome.
6) Select Show Signal and / or Show Envelope for reference.
7) Use the metronome sliders to adjust the displayed metronome beats (optional).
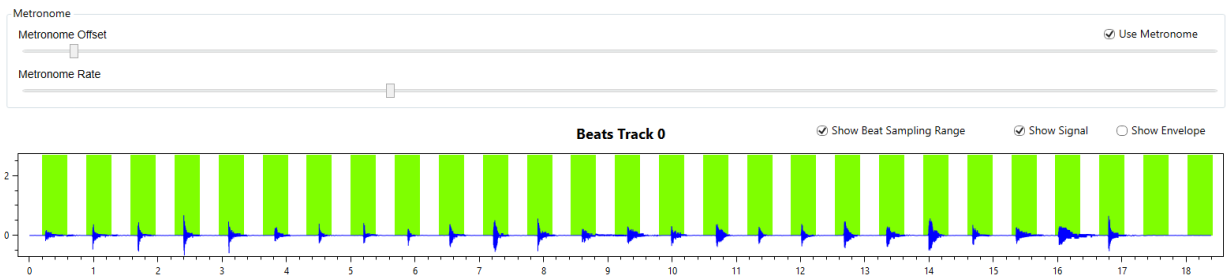8) Set Score Options and click Score.

The Combine Results selection will determine how transcription is performed.

Use FFT Profile Results Only will use the FFT selected in the Octave Mapping expander and apply the profiling method selected in the option dialog.

Use Spectrogram Analysis Results Only will transcribe from the Spectrogram results; either by direct mapping from the Constant Q spectrogram, or by applying the selected Convolutional Neural Network to the spectrogram used for training.

Selecting the instrument to score will load and display the model provided for the instrument if it is available. The user may also select and load a CNN that they have produced using the CNN tab.

Once the beats are defined the score options may be selected and the Score button clicked. The score will appear in the score tab.
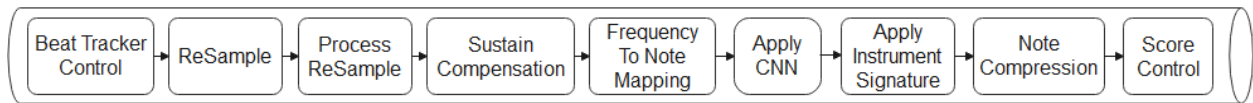
The Score Workflow

Figure 12 : Score Workflow

The Score Workflow is initiated when the Score button is clicked. For each detected beat of metronome beat the Beat Tracker Control will send a beat event to the ReSample task.

The ReSample task will resample the FFT size selected in the Octave Mapping expander starting from the beat time. This ensures that the FFT size is sufficient to map all required notes and is accurately timed to the beat.

The Process ReSample task will re-calculate the FFT for the re sampled data.

The Sustain Compensation task will perform sustain compensation if this is selected in the options dialog. This will help to prevent notes from bleeding over into the next sample. The amount of compensation is variable from 0 – 0.7.

The Frequency to Note Mapping task will calculate the energy for each note in the allowed octaves.

The Apply CNN task will apply the selected Convolutional Neural Net model to the spectrogram.

The Apply Instrument Signature task will apply the instrument profile if one exists for the selected instrument to score. It will combine the profile results with the CNN results from the Apply CNN task; then it will create the actual notes for each beat.

The Note Compression task will calculate the length of each note / cord and create the bars and stave information. As each stave is completed it will be sent to the Score Control where it will be displayed.

CNN Tab
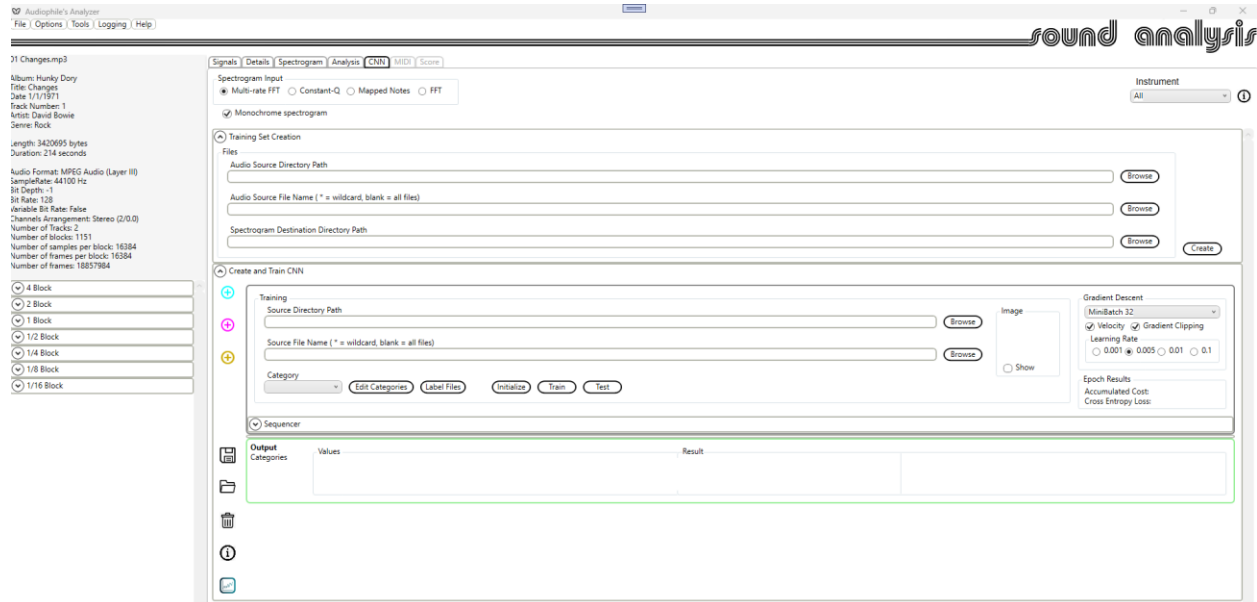
This tab provides the tools needed to configure and train a Convolutional Neural Network to use for transcription. This is not required if the user does not wish to use a CNN for transcription or will use one of the provided CNN models.

The full functionality of this tab is described in a separate manual:
 "Audiophile's Analyzer 2.0 Convolutional Neural Network Tab"

MIDI Tab
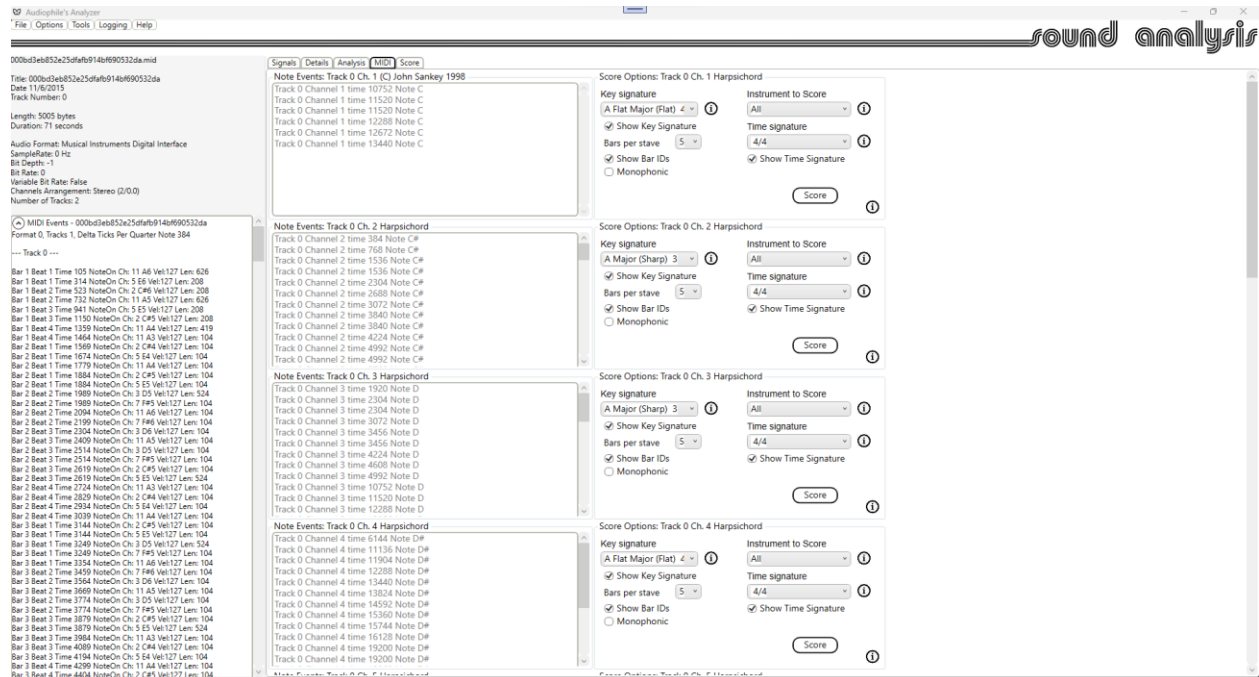
Figure 13 : MIDI tab

When a MIDI file is opened the MIDI tab will open. The MIDI events will be shown by track in a left column expander. For each channel a note events control and a score options control will be displayed. The note events panel will list the notes for a channel that will be transcribed. The score options control will allow the user to change the options and score the channel. The transcribed score will appear in the Score tab.

Figure 14 : MIDI Workflow

Score Tab

The score tab will show the transcribed score once the Score button is clicked.
The Save to MIDI button will save the displayed score as a MIDI file. The save MIDI events will be displayed in a left column expander.
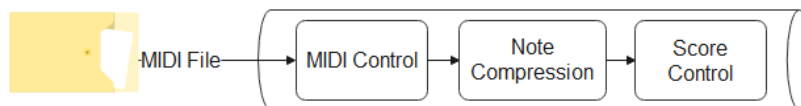

Figure 15 : Score Tab

Options
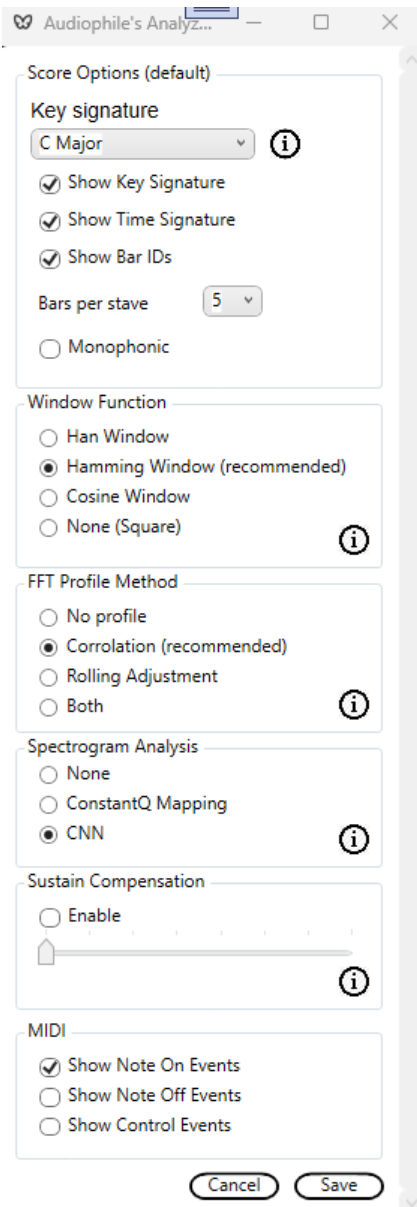
Figure 16 : Options Dialog

The options dialog can be launched from the menu bar.

The score options for the session are defined in the analysis tab. The defaults are defined in the options dialog. These options are concerned with the score display and do not affect the transcription process. The default key signature is only applied when a key signature cannot be calculated from the FFT totals.

The Window function will change the way that the audio samples are windowed before the FFT is applied. A full description of window functions can be found at http://en.wikipedia.org/wiki/Window_function

The profile method affects the way that the profile (instrument signature) is implemented. Correlation is recommended for single instrument files. Turning off the profile is a good way to assess the profiles effectiveness.

Spectrogram analysis can be either Constant Q mapping where notes are directly mapped to the Constant Q spectrogram, or CNN where a convolutional neural network can be applied to any spectrogram (usually the multi-rate FFT spectrogram).

Sustain compensation will help to prevent notes from bleeding over into the next sample. The amount of compensation is variable from 0 – 0.7.

MIDI options will select which events are displayed in the MIDI control.

Instruments

Instrument selection dropdowns are used in the analysis tab to select the instrument to use for rhythm tracking, and the instrument to score. A number of instrument definitions are built into the Audiophile's Analyzer; these definitions may be edited, or others added.

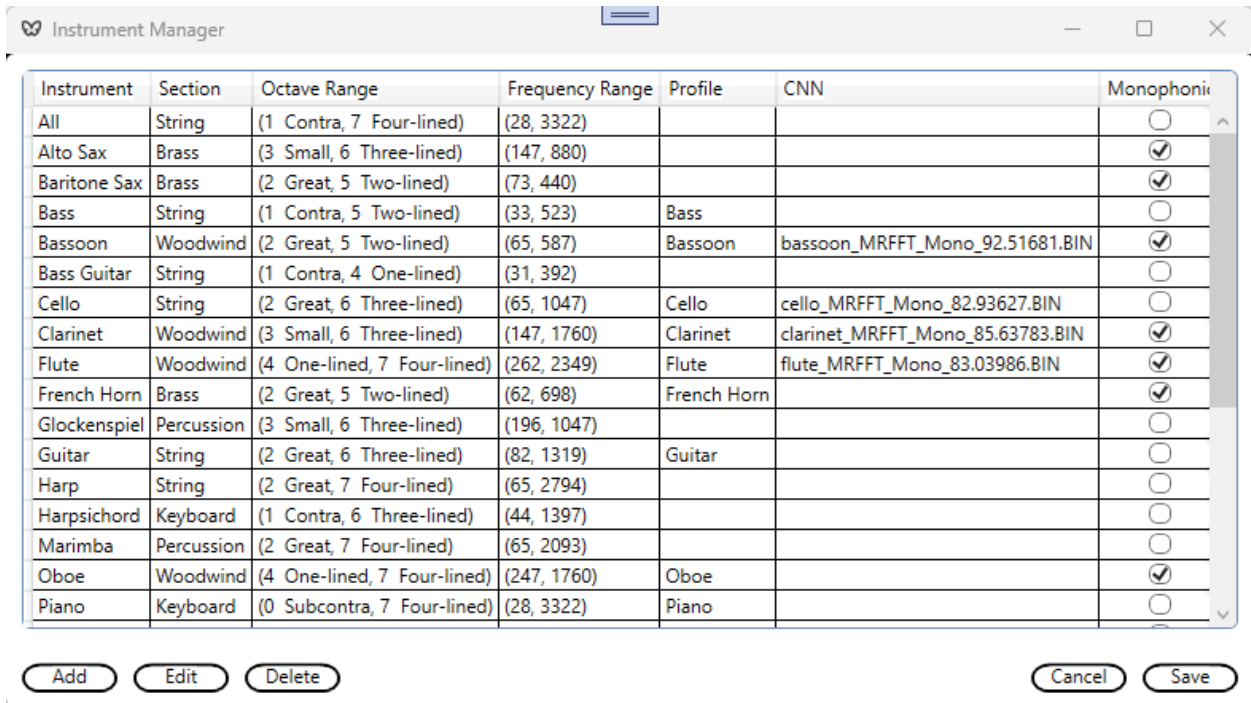The instruments manager can be launched from the Tools menu.

Instrument	Section	Octave Range	Frequency Range	Profile	CNN	Monophonic
All	String	(1 Contra, 7 Four-lined)	(28, 3322)			○
Alto Sax	Brass	(3 Small, 6 Three-lined)	(147, 880)			✔
Baritone Sax	Brass	(2 Great, 5 Two-lined)	(73, 440)			✔
Bass	String	(1 Contra, 5 Two-lined)	(33, 523)	Bass		○
Bassoon	Woodwind	(2 Great, 5 Two-lined)	(65, 587)	Bassoon	bassoon_MRFFT_Mono_92.51681.BIN	✔
Bass Guitar	String	(1 Contra, 4 One-lined)	(31, 392)			○
Cello	String	(2 Great, 6 Three-lined)	(65, 1047)	Cello	cello_MRFFT_Mono_82.93627.BIN	○
Clarinet	Woodwind	(3 Small, 6 Three-lined)	(147, 1760)	Clarinet	clarinet_MRFFT_Mono_85.63783.BIN	✔
Flute	Woodwind	(4 One-lined, 7 Four-lined)	(262, 2349)	Flute	flute_MRFFT_Mono_83.03986.BIN	✔
French Horn	Brass	(2 Great, 5 Two-lined)	(62, 698)	French Horn		✔
Glockenspiel	Percussion	(3 Small, 6 Three-lined)	(196, 1047)			○
Guitar	String	(2 Great, 6 Three-lined)	(82, 1319)	Guitar		○
Harp	String	(2 Great, 7 Four-lined)	(65, 2794)			○
Harpsichord	Keyboard	(1 Contra, 6 Three-lined)	(44, 1397)			○
Marimba	Percussion	(2 Great, 7 Four-lined)	(65, 2093)			○
Oboe	Woodwind	(4 One-lined, 7 Four-lined)	(247, 1760)	Oboe		✔
Piano	Keyboard	(0 Subcontra, 7 Four-lined)	(28, 3322)	Piano		○

Add Edit Delete   Cancel Save
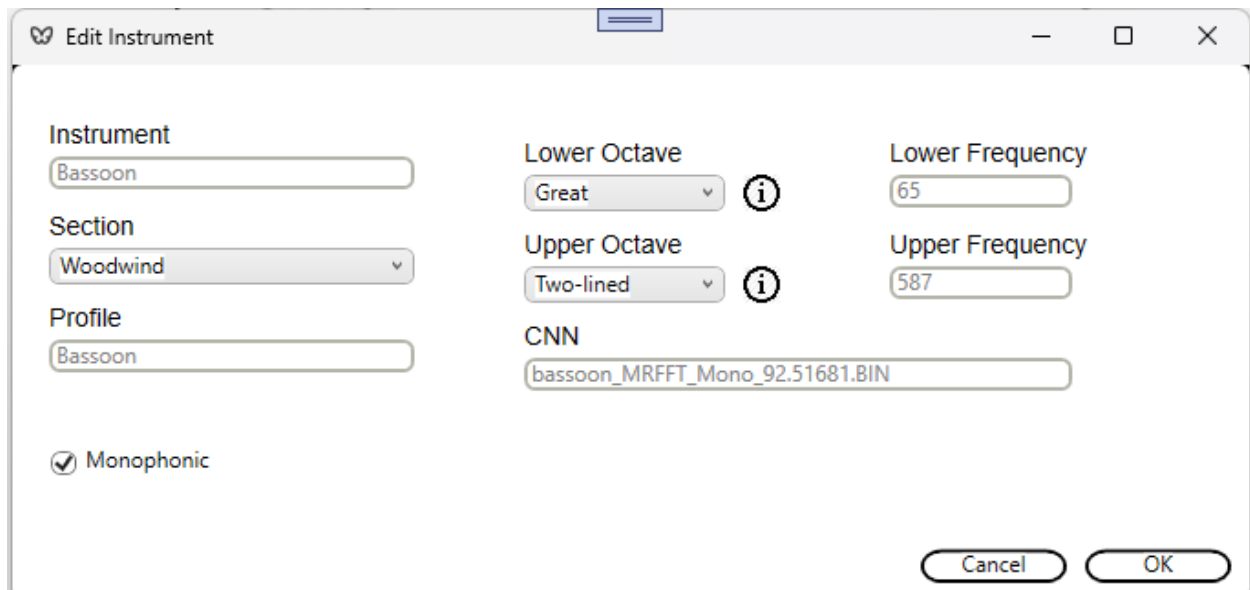
Figure 17 : Instruments Dialog

Figure 18 : Edit Instrument Dialog

Profiles

A profile can be assigned to an instrument. This is used to handle the problem of overtones. The profile will model the overtone pattern of a single note from the given instrument.

Figure 19 : Profile Manager Dialog

To create a profile:

1. Open the profile manager.
2. Click the green button which turns red.
3. Open a file with a single note from the required instrument with the main window of the Audiophile's Analyzer.
4. Select the octave and note from the dropdowns.
5. Use the analysis tab to score the note.
6. Repeat steps 3 – 5 multiple times with different sample files.
7. Click the red button to stop sample collection.
8. You will see a list of collected samples which you can display by selecting them. They should all look very similar.
9. Work down the list and un-check any samples which look different from the pack.
10. Once you are satisfied with the selected samples click 'Create Profile'.
11. The created profile will be displayed.
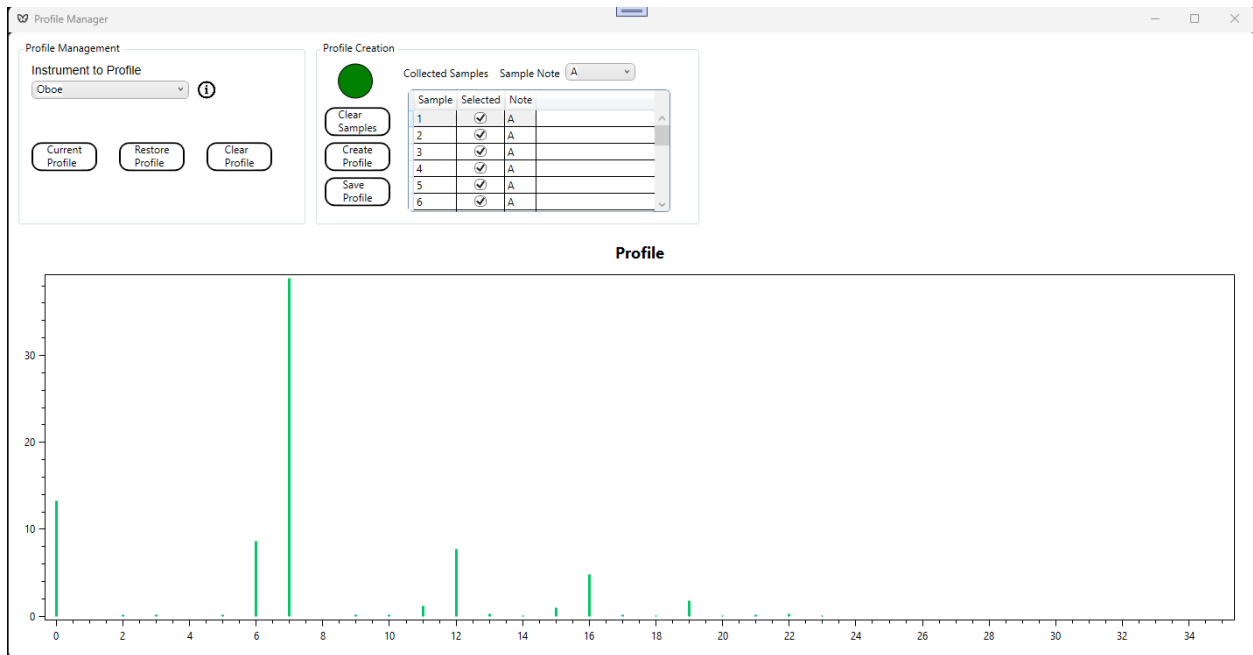12. To save it; click Profile -> Save Profile from the menu bar.

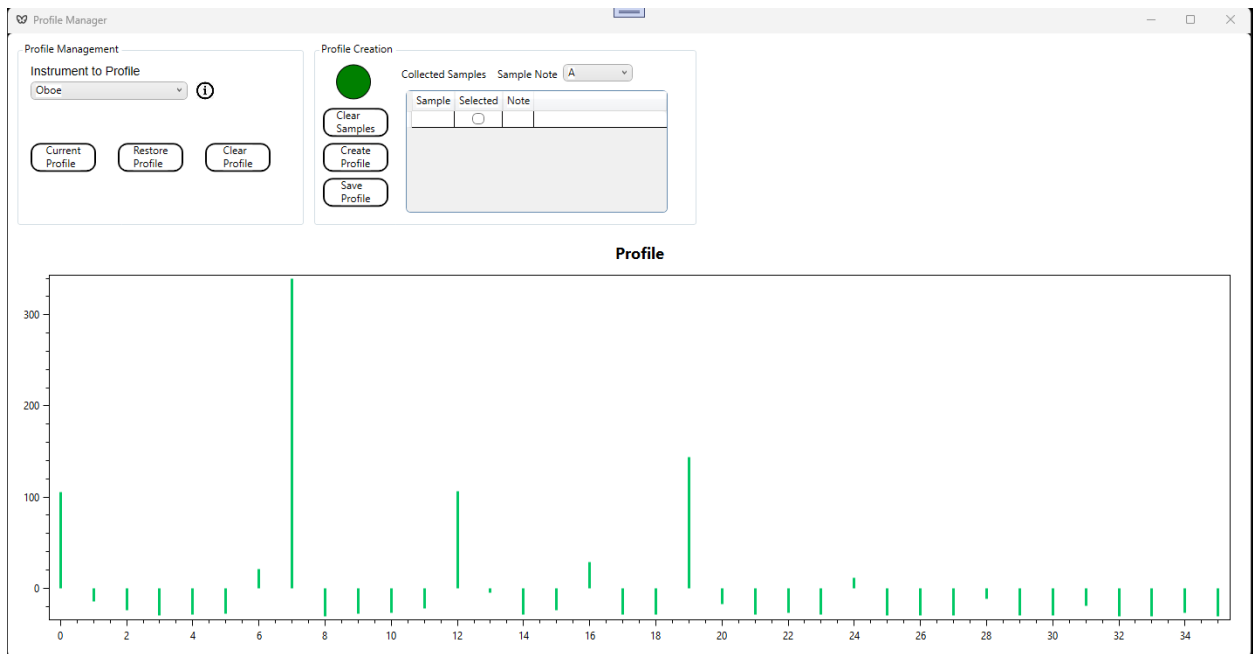**Figure 20 : Profile Creation: step 8 list of collected samples**

Figure 21 : Profile Creation: step 11 the created profile

The advantage of creating a profile

We have just seen the creation process for the oboe profile. The oboe has a very strong first harmonic.
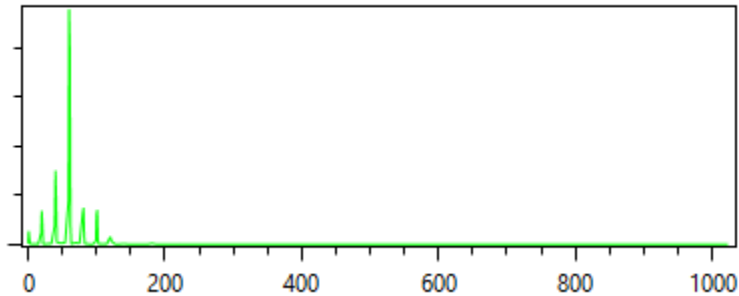
By 4 (0)

Figure 22 : Oboe spectrum

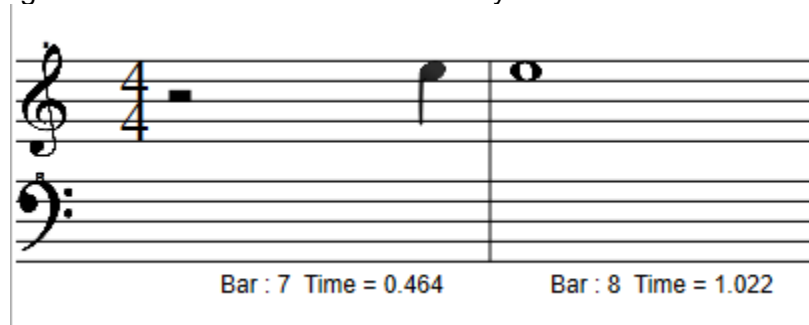Without profiling the second harmonic will incorrectly score as E.

Bar : 7 Time = 0.464   Bar : 8 Time = 1.022

Figure 23 : Oboe note A without profile

With profiling the note is correctly scored as A.

Bar : 6 Time = 0.650   Bar : 7 Time = 1.022

Figure 24 : Oboe note A with profile

Sound Analysis   **24**

The Score Editor

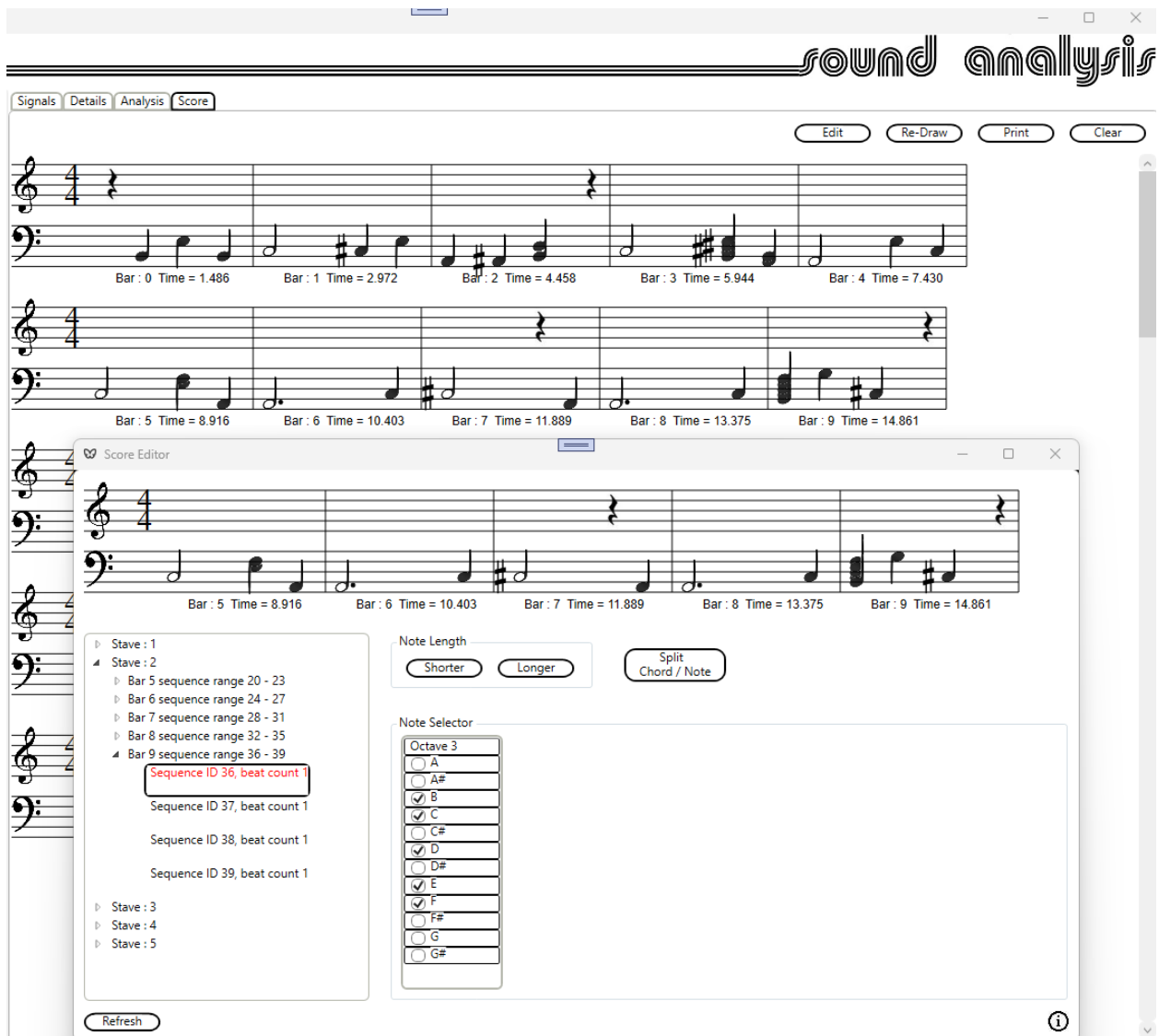The Score Editor tool is designed to allow the user to easily make quick fixes to the transcribed score.

To launch the editor, click the Edit button in the score tab.

The editor provides a tree view of the current score. Expand this to the stave, bar and chord/note/rest. Double click the chord/note/rest (Sequence Id) of interest. The selected stave will be displayed; the notes in the chord/note/rest will be selected in the note selector. Use the note selector to change the notes in the chord/note/rest; the displayed stave will update automatically. The longer and shorter buttons can be used to change the length of the chord/note/rest. The number of beats to the bar will be maintained. The following chord/note/rest will be adjusted to maintain the beats to the bar; if necessary, an extra rest will

be added. The Split Chord/Note button will divide the selected chord/note/rest into two. The tree view will update automatically.

Re-Draw the score to update it with the current edits, while the editor is open or after it is closed.

Changing the score will not affect the rest of the session data, samples, etc. So, re-scoring the session will ignore the edited changes.

Help Menu

Show Users Guide will launch this document.

Show CNN will launch the manual for the CNN tab.

About will show a dialog with the version and other information.

The info symbol ⓘ will appear in many places, clicking this will bring up a help message about the control associated with the symbol.

Performance

Since the input is from a file there is no requirement for real-time operation.

Transcription accuracy will vary depending on the file in question.

The simplest files are single instrument files. With correct profiling these can be transcribed with reasonable accuracy. Note that a profile is not simply for a specific instrument, but may apply to an instrument and style of playing.

Files with multiple instruments are more challenging. An instrument is defined by its frequency range, allowed octaves, and profile. Using the correlation method for profiling is recommended. Where an instrument can be separated easily by frequency range, somewhat reasonable results may be obtained.

Separating instruments from a full orchestra is not feasible with existing technology.

Contact

If you have any questions, suggestions, or feedback email steve@sound-analysis.com

To report a bug please include as much useful information as possible; for example, a short description of the problem, the audio file used, open the debug log viewer and copy/paste the contents to the email, or a screen shot, would all be useful.