

The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals

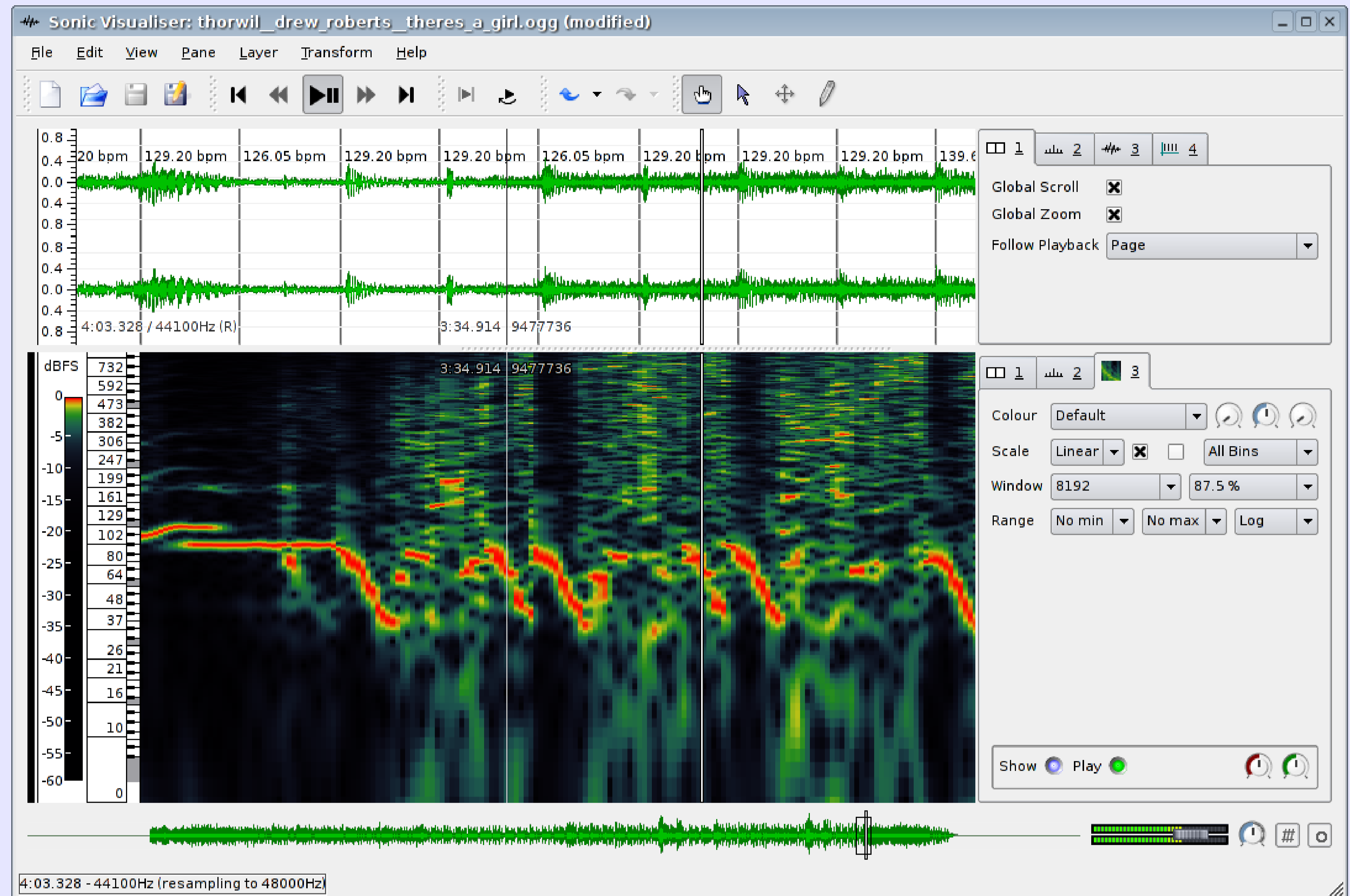
Chris Cannam, Christian Landone, Mark Sandler, Juan Pablo Bello
Centre for Digital Music, Department of Electronic Engineering, Queen Mary University of London

The Sonic Visualiser is an application for the automatic extraction, browsing and editing of time-synchronous audio descriptors and is designed to assist the study and comprehension of the contents of audio data, with particular emphasis on musical recordings.

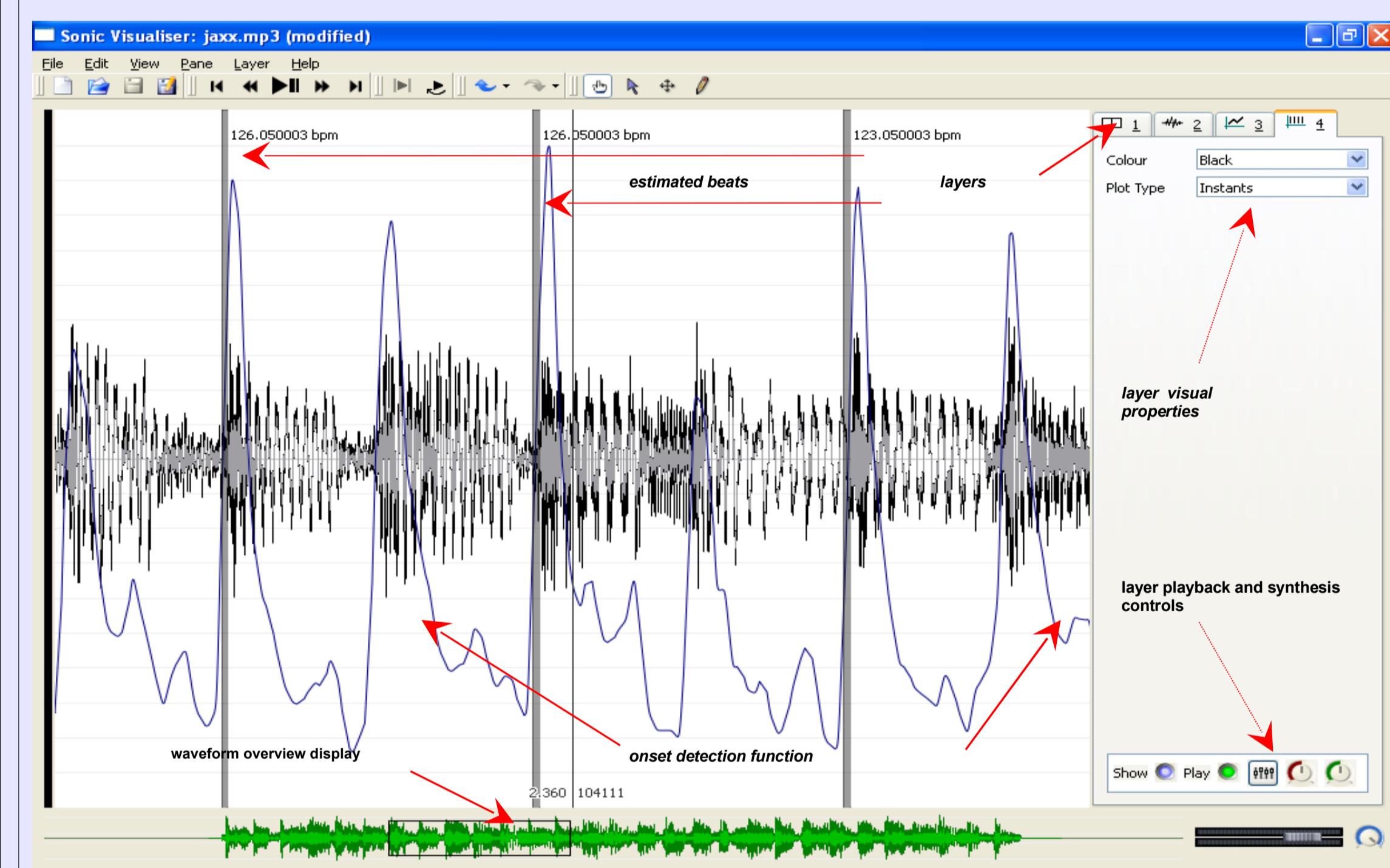
The architecture of the application allows for easy integration of third party algorithms for the extraction of low and mid-level features from musical audio data.

Highlights:

- > Cross-Platform (Linux, MacOSX, Windows)
- > Released under the GNU General Public License
- > Automated feature extractors available as plugins
- > Time-synchronous features are displayed as editable and overlaid logical layers
- > Available for download from : www.sonicvisualiser.org



Overlaid Data Visualisation



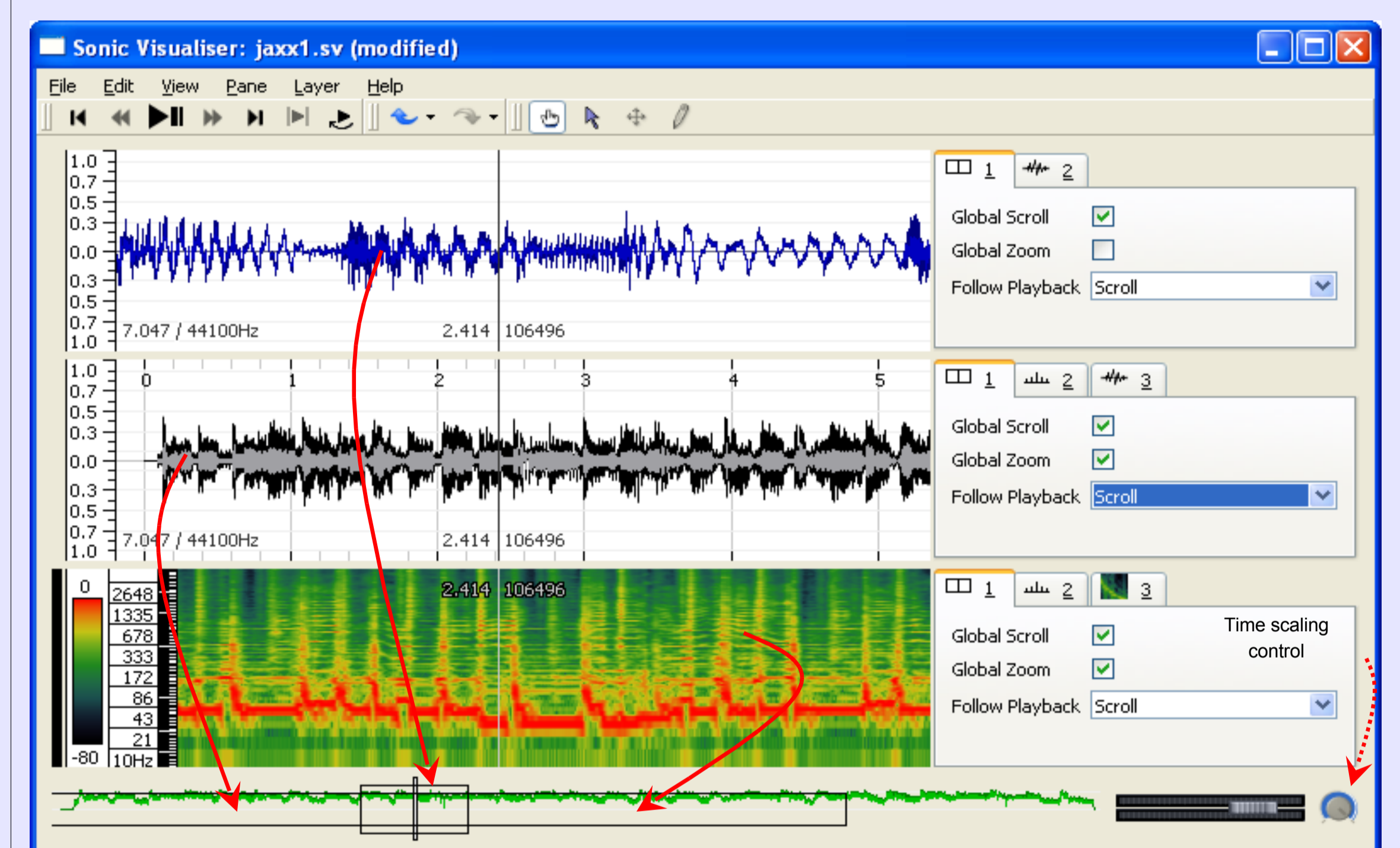
The x-axes of separate layers on the same pane always align exactly by audio frame and a layer can be visually raised above the others by selecting its property box from the stack. The visual properties of each layer, such as plot type, colour, scale and magnification, can be modified by the user

The Sonic Visualiser uses stackable panes that can be added or removed at the user's request and can show a number of different pieces of data overlaid on one another and aligned according to time.

The overlay approach greatly assists comparisons among extracted and annotated features, as well as aiding understanding of the original audio data by allowing it to be viewed in more than one form "at once". This facility may be used with various sorts of data, including points in time, values on a discrete time graph, as well as waveforms and spectra.

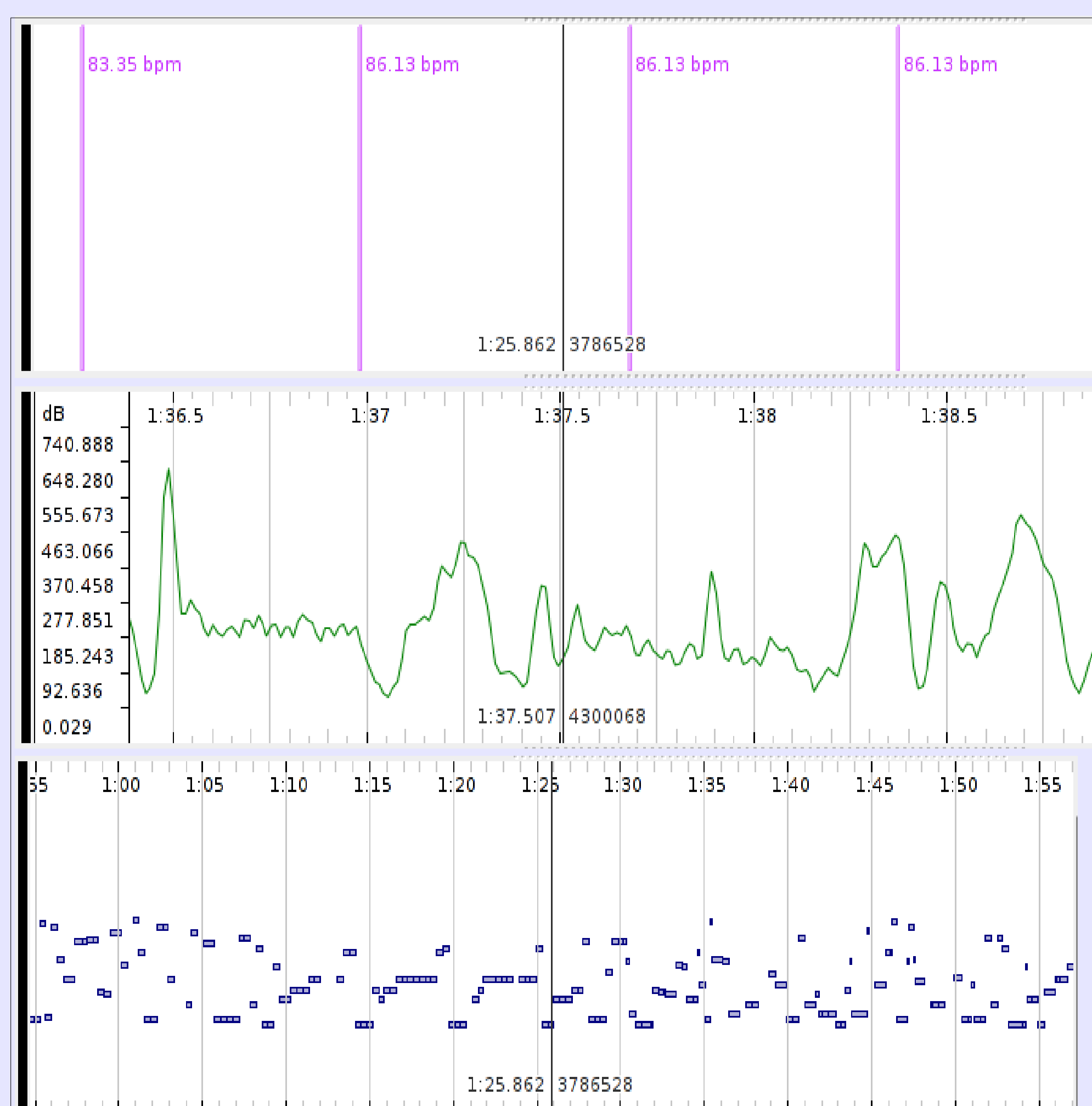
In addition to the user-definable panes, an overview of the time domain waveform under analysis is shown in a separate area at the bottom of the application's user interface.

Time-Aligned Stacked Panes



The panes are usually aligned to the same sample frame and the user can scroll through the audio by dragging the pane with the mouse. However, the separate panes do not have to be shown at the same time resolution and different zoom levels can be applied independently, allowing a multi-resolution representation of the same data.

Annotation Layers



The Sonic Visualiser provides the user with different types of manual annotation layers that can be exported to a file or used to import and visualise comma and space separated, or MIDI data files.

The "Time Instants Layer", shown at the top, allows the insertion of vertical lines at points in time that may represent onsets or beats, while the "Time Values Layer" (middle) is an editable sequence of points in time, where each point has a value that determines its position on the Y axis. Each point may also have a label.

The notes layer (bottom) consists of an editable sequence of notes with an associated label where each note has a start time, a duration and a value that determines its position on the Y axis. A text layer is also available for placing freely around the pane informal or descriptive annotations that are not specifically attached to other features.

The Sonic Visualiser allows entire work sessions to be saved to file, enabling the instant recall of the audio file under analysis, visualisation styles as well as settings, annotations and data layers.

Playback of audio and features

Audio data can be played by the Sonic Visualiser using the provided transport commands.

Looping and time stretching can be used during playback to aid analysis and annotation and the amount of time scaling can be varied in real time using a rotary control in the user interface.

Features can also be "listened to": each of the data layers displayed in the Visualiser has a set of performance characteristics, including auralisation method (e.g. play a particular sound each time a data point is hit during playback), output level, and stereo pan.

It is therefore possible to compare aurally a set of calculated or annotated features by panning them separately or playing them with different sounds.

Vamp plugins

An external API is used by the Sonic Visualiser for the automatic extraction of low and mid-level descriptors from musical audio data.

Vamp plugins are dynamic link libraries that process sampled audio data, returning complex multidimensional data with labels representing semantic observations that are automatically loaded by the Sonic Visualiser and shown on a new layer.

The cross-platform plugin SDK is distributed under the terms of a BSD-style license, allowing closed-source feature extraction modules to be used by the Sonic Visualiser and thus permitting the safe dissemination of musical analysis algorithms that are yet unpublished and/or not protected by patents