

Sonic Visualiser Report

About Sonic Visualiser

Sonic Visualiser is an open-source application for editing audio files. Sonic Visualiser was created and developed by the Centre for Digital Music at Queen Mary, University of London in 2007, and the most recent version was released in December 2019. For the purposes of this project, Sonic Visualiser is being tested and evaluated as a tool for annotating audio files using the mark-up tools within the application.

Annotation Capabilities in Sonic Visualiser

Layers

Sonic Visualiser's main editing interface involves panes and layers. The pane is the horizontal canvas where the waveform of an audio file can be viewed. This is the main navigation area of the audio file, and is horizontally scrollable. The waveform portion of the pane is a layer. Layers on Sonic Visualiser are components that can be added and edited and are stacked on top of each other. Multiple layers can be added and stacked. Multiple panes can be added as well.

The following layers were evaluated for their potential to be useful as a way of annotating audio recording for the purposes of this project:

- **Text Layer:**

The text layer allows for text to be added horizontally and vertically across the waveform in the audio file. These text points have a data point across time and vertically in the waveform. Data from the text can be exported. This layer works best for adding visible annotations to an audio file.

One helpful aspect of adding annotations via the text layer is that points can be organized in different colors, so as an annotator or archived audio this is visually straightforward and could help with organizing and categorizing annotations. The downside is that only one color can be added per layer, so multiple annotation layers would need to be added. These layer data files could be added and combined manually. Naming a layer essentially creates a template that can be reused later.

In a use-case example, a recording of "Her Kind," was opened in Sonic Visualiser, and annotations were added in four categories: repetition, imagery, PoemStructure, and paratext. A layer was added for each category to facilitate simple and organized annotation creation. Points were added throughout the poem: "Her Kind," a repeated phrase was added in the repetition category; bodyreference was a tag added to reflect

January 13, 2020

any mention of the body in the imagery category; stanza beginning and ends as well as title points were added to PoemStructure; and breath was added where audible breathing could be heard, in the paratext category. Each individual layer was then downloaded, and the results were able to be combined into one file.

- **Time Instants Layer**

Another layer is the time instants layer. The points on this layer represent a moment in time and have a layer on the upper corner. Adding a time instants point is extremely user friendly: while the audio is playing, all you do is click the semicolon. A point is then automatically added. The user can click and drag to move around to create a more precise point, and the label can be added by entering the table data. The main issue with the Time instants layer is that the text label is only visible in the data table and exported data. The visual label within the user pane is the time stop. But, this method of adding points is quite intuitive, and would be more conducive to archival annotation if a text label was easier to add here.

- **Regions Layer**

The regions layer is used to mark segments of ranges across time. This feature is not intuitive and is difficult to make precise. Once a range is added, the range cannot be manipulated/stretched in length/direction, but rather, a click and drag will move the location of the entire segment. This layer would be difficult to integrate into archival annotation settings. If the layer's functions could be improved to include a more easily editable region range that could be dragged to edit the duration, this could be a helpful feature.

Exporting data

Data added in annotations can be easily exported into a .csv, .txt, .ttl, or a .svl file. Text layers are exported with data corresponding with time, frame, and height. Time instants layers are exported with columns for time, frame, and label. Regions layers are exported with columns for time of the point, frame, value (which defaults to 1), duration, and a label. The downside to exporting data is that layers have to be downloaded individually. It would be advantageous to have a way to export multiple layers at once. The workaround for this is to edit in one layer, or to edit in the same kind of layer.

Importing data

Data can be imported into Sonic Visualiser from a .csv or .txt file. For example, if you wanted to import labels you had created on the same audio in Audacity, you can download the same labels in Sonic Visualiser by importing the file that includes time data into a layer. To do this, you simply navigate to a new time instants layer. Then, under the File dropdown, you click,

January 13, 2020

import annotations layer, then upload the file and select the column types from the dropdown. The problem with this is, it seems to only want to upload as a time instants layer. It would be helpful if you could choose the type of layer to upload. This also won't work with a mix of annotations and ranges, unless you ignore the end time of the range column.

Overall Evaluation

Sonic Visualiser is acceptable for adding annotations. At the same time, many of the methods for annotating are limited. Further, the application itself is not very intuitive and makes point manipulation and precision along the waveform difficult at times, especially with the regions layer, which could be a very helpful tool in annotating archival audio. One helpful feature when it comes to editing is the ability to open up the data table of a layer at anytime and manipulate points and labels (this is done by navigating to Layer > Edit Layer Data or pressing "E" at any time).

Another positive element of Sonic Visualiser are the plugins the application supports. Sonic Visualiser has a suite of plugins that have been developed to extend the capabilities of the application's features. While the majority of these plugins are geared towards musicology and music annotation, some of them have potential for further understanding and critically thinking about archival audio, even when music is not present. For example, one plug in, called Segmentation by Voice/Music, allowed allows segmentations to be created based on where speech and music are detected.

It would be helpful to have a way to merge layers when exporting files, as the only current option is exporting individual layers. Another missing element of Sonic Visualiser is the ability to shortcut adding annotations in the Text layer. This is possible in the Time Instants layer, by pressing the semicolon key at any time. The Text layer is the best option for annotations, but it's cumbersome to add a point precisely. It would be helpful if there was a key shortcut for placing a point while listening to the audio. An easier way to add precise regions layers is another missing element. Right now, a region can be added, but the range (the time covered in the region) is not possible to adjust. The entire section can be moved using the move tool, but the range cannot be edited by adjusting at either end of the region points.